HZ BOOKS
华章心理

语言风格的秘密

语言如何透露人们的性格、情感和社交关系



The Secret Life of Pronouns

What Our Words Say About Us

[美] 詹姆斯·彭尼贝克（James W. Pennebaker）著 刘珊 译

著名社会心理学家/语言学家集20年研究之大成

教你如何通过分析语言风格了解自己和他人



机械工业出版社
China Machine Press

VISIT...

LANZAROTE Caliente.com


华章心理

语言风格的秘密

语言如何透露人们的性格、情感和社交关系



The Secret Life of Pronouns

What Our Words Say About Us

[美] 詹姆斯·彭尼贝克（James W. Pennebaker）著 刘珊 译

著名社会心理学家/语言学家
集20年研究之大成

教你如何通过分析语言风格
了解自己和他人


机械工业出版社
China Machine Press

语言风格的秘密：语言如何透露人们的性格、情感和社交关系

The Secret Life of Pronouns：What Our Words Say about Us

（美）詹姆斯·彭尼贝克（James W. Pennebaker） 著

刘珊 译

ISBN：978-7-111-58415-5

本书纸版由机械工业出版社于2017年出版，电子版由华章分社（北京华章图文信息有限公司，北京奥维博世图书发行有限公司）在中华人民共和国境内（不包括香港、澳门特别行政区及台湾地区）制作与发行。

版权所有，侵权必究

客服热线：+ 86-10-68995265

客服信箱：service@bbbvip.com

官方网址：www.hzmedia.com.cn

新浪微博 @华章数媒

微信公众号 华章电子书（微信号：hzebook）

目录

前言
第1章　最容易被人遗忘的词汇的秘密
　当社会心理学和语言学相遇
　通过计算机程序识别健康的写作
　语言像镜子，也是工具
第2章　重要的不是内容，而是语言风格
　用词风格是心理的钥匙
　语言风格和心理：跨向下一阶段
第3章　语言体现性别、年龄和权力
　性别、刻板印象和计算机数据
　两性间的语言差异有多大
　对人物性别的把握：剧作家和编剧
　语言中的性别差异：男性荷尔蒙的作用
　语言的智慧：一生中使用的语言
　社会阶层和语言
　谈到性别、年龄和社会阶层：权力之声
第4章　性格：发现人的内在
　通过功能词了解人的性格
　通过图片发现性格
　通过语言了解他人和我们自己
第5章　情绪测定
　不同的情绪，不同的思考方式
　代词和痛苦：自杀性诗人的语言
　傲慢、失落和沮丧：朱利安尼市长和李尔王
　创伤如何显露：语言是窗口
　追踪一个国家的情绪：2001年9月11日的博客
　“9·11”事件之外：天灾和人祸
　使用语言分析来指导情绪波动后的心理健康治疗
　创伤的长期影响：语言如何治愈
　将秘密转化为语言：表达性写作
　两者结合：情感和思考方式是硬币的两面
第6章　说谎的语言
　自我欺骗的语言
　谎言、骗子和无赖的语言：寻找语言测谎器
　事关紧要时捕捉欺骗：避免入狱、心脏病和战争
　监察：哪些词汇是真实和欺骗最好的预言者
第7章　语言体现地位、权力和领导力
　弄清楚谁是头儿：解码社会层级

通过通信发觉自己的地位
陌生人对话中社会层级的快速显现
定位尼克松当局领导人的地位
领导阶层的语言
有关地位、权力、傲慢和领导力的最后思考
第8章　爱的语言
语言模仿
语言风格匹配和大脑
体验对话中的“LSM过山车”
创造一个爱的语言检测器
通过LSM了解过去的关系
通过LSM了解亲密关系
第9章　通过语言了解团体、公司和社区
“我们”作为身份标识
扩展的“我们”：人和团体
团体认同之外：捕捉团体动态
团体定位：语言如何显示一个团体在做什么以及它在哪里
第10章　语言侦察
通过语言识别作者
谁写的？作者的身份识别
歌词揭示乐队情况：披头士乐队
小结：整理你的作者识别工具包
政治事件和历史事件的语言线索
这些人是谁？尝试了解美国总统
“我”说了什么？—奥巴马少用的人称代词
因攻击和战争避免使用“我”
通过语言重看历史
通过人们的语言推测他们未来的行为
通过大学录取申请书的语言预测大学成绩
通过语言预测出狱后的生活
附录　识别和解释功能词的便捷指南
注释
参考文献

前言

回想一下你的最后一次谈话、你发出的最后一封邮件或最后一条短信。你以为你在谈论的是晚饭吃什么，你要洗的衣服，下一次销售会议的策略。也许你确实谈到了这些。但与此同时，却不只是这些，你说了更多的内容。你在交流信息时使用的词汇比想象中揭示了更多关于你的情况。你、一个、是、到、我、但是、这、为、不……

代词、冠词、介词和其他大量的小词、隐性词揭露了你的个性、思考方式、情感状态和与他人的沟通状态。这些词通常被称为功能词，虽然在词汇表中占比不到1/1000，但在你日常使用的词汇中占比却将近60%。你的大脑可能没有注意到这些词，但是如果仔细观察，你将会发现它们微妙的力量。

功能词的作用可能与你想象中的不同。举个例子，“我”，是英语口语中最常用到的词，领导人使用它的频率比追随者使用它的频率高出很多，说真话的人使用它的频率比说谎的人高出很多。使用冠词频率高的人[某个、这个（a、an、the）]比使用冠词频率低的人在大学的表现好。还有，如果你想找到你的真爱，可以和你的准伴侣对比你们使用功能词的方式。

虽然本书关注的是功能词，但是它确实不是关于讲话的。相反，本书讲的是如何通过这些词来探究人的性格和社会关系。我的学习和早期的研究建立了社会、性格、临床和健康心理学间的桥梁。通过一些偶然的发现，我才注意到了这些隐性词汇的存在。

刚开始，对功能词的研究只是尝试。但是随着更加深入的研究，我开始觉出它们与领导力、心理健康、脑功能等问题有着意想不到的联系。紧接着，我和我的学生们就花时间与电脑工程师、语言学家、联邦调查局特工、律师、医生和营销大师在一起，还有我的同事们，有教历史的、政治学的、通信的，甚至还有教会计学的。而且，最近，我们被社会媒体的狂热所席卷——我们将我们的想法和方法应用到推特、脸书、线上约会、博客、短信、邮件，甚至是过时的电话聊天中。

本书有些像旅游指南。它是围绕着几个我最感兴趣的有关心理和社会科学的话题开展的，如性格、性别、欺骗、爱、历史、政治和团体。目的是展示如何通过分析功能词来对每一个话题提出新的见解。与此同时，我想让大家领会如何思考和分析语言。不管你的个人兴趣或专业兴趣是什么，我希望你可以用不同的方式来看这个世界，并可以利用这些知识来更好地了解自己和他人。

***

语言分析正在经历一场革命，这将会对社会科学和人文科学产生深远影响。人的一生使用的语言就像是指纹。渐渐地，这些语言可以用来确定他们的身份，甚至是他们的背景。语言的使用，尤其是功能词的使用，可以显示一个人的社会关系网以及他们在家庭、邻里和工作中扮演的角色。只要有一点语言痕迹，就有许多基于计算机的方法可以捕捉到。莎士比亚、孔子、古老的宗教经文、政治家和小说家留下的文字数量令人感到惊异，各个学科的学者可以用新的角度和前所未有的工具来分析这些文字。

虽然语言分析是本书的重点，但它是一本心理学书籍。语言学家感兴趣的主要是语言本身，而我关心的是语言所反映出的人们的心理状态。语言可以被认为是洞察人们的思想、感受、动机和与他人联系的有力工具。随着计算机技术的进步，这种方法正在被各个学科的学者所了解，如语言学、社会语言学、人类语言学、神经系统科学、心理语言学、发展主义、计算机科学、计算机语言学以及英语学者和其他学科的学者。一些最富创新性的工作来自形形色色的学者和公司之间的合作，比如谷歌。

如果你对本书中的话题和方法感兴趣的话，有大量由睿智的思考者和作家创作的作品是面向广大读者的。对我的方法影响最深的有社会学家欧文·曼欧文、语言学家乔治·莱考夫、认知科学家史蒂芬·平克、社会语言学家黛博拉·坦纳和人类学家安娜·威尔兹彼卡。

其他人的帮助也直接影响了我对词汇使用和心理状态之间的联系的思考。弗洛伊德早期对口误的研究标志着第一个心理学认知，即突然冒出的无意识话语可以揭示人们隐藏的想法和感受。后来的心理分析学家，如路易斯·戈特沙尔克和沃尔特·温特劳布，为如何将看似无关痛痒的话与更深层次的动机和病人表达的担忧联系在一起提供了路线图。

几个基于计算机的文本分析系统预先显示了我需要的计算机程序，尤其是菲利普·斯通20世纪60年代开发的通用询问程序。任何关于文本分析初期的故事也要提到以下名字，如道格·拜伯、赫伯·克拉克、唐纳德·福斯特、霍华德·贾尔斯、罗德·哈特、罗伯特·霍根拉德、汉斯·科尔迪、克劳斯·克里彭多夫、科林·马丁代尔、艾哈德·默根瑟勒和其他一些人。

然而，这些年，影响我最深的人是我的学生和同事。本书集合了几十个甚至是上百个人的努力。当前语言项目的核心人物包括詹娜·巴德利、大卫·比弗、罗杰·布思、玛莎·弗朗西斯、阿特·格雷泽、卡拉·格鲁姆、杰夫·汉考克、莫莉·艾尔兰、埃娃·凯斯维兹、劳拉·金、马蒂亚斯·梅尔、凯特·尼德霍夫、基思·皮特里，奈兰·拉米雷斯-埃斯帕扎、斯蒂芬妮·鲁德、伊泰·赛、理查德·斯拉切、洛瑞·斯通·汉德尔曼和依拉·陶斯兹克。比尔·斯旺、鲍勃·约

瑟夫，以及在得克萨斯大学奥斯汀分校心理学系的其他同事在指导我的工作方面也起着很重要的角色。

没有国家科学基金会、陆军研究所、美国国立卫生研究院和德克萨斯大学的人文科学学院的资助，就不会有这本书。我特别感谢书语者对本书的支持，马克·阿伦森、迪尔德丽·穆莱恩和布鲁姆斯伯里出版社的编辑们，彼得·金纳和皮特·比蒂。

有三个人对本书的完成影响最大。辛迪·钟多年来一直陪我处在研究工作的核心位置。她的想法、反馈和幽默使得研究旅程非常愉快。我的同事和研究伙伴山姆·高斯林要求我更全面地思考有关个性和社会进程的问题。最后，我的妻子露丝一路上鼓励和指导我，并督促我用新视角去思考故事、社会关系和真实世界。她的精心指导体现在本书的每一页上。

第1章　最容易被人遗忘的词汇的秘密

那些最能透露信息的词是最短的和最容易被人遗忘的词。代词、冠词、介词和其他隐性词汇能透露出我们是谁。

大家早上好！希望你的一天都精彩！亲亲帕丽斯:)

——帕丽斯·希尔顿，媒体人物

昨天在贝鲁特会见了瓦立德·琼布拉特——德鲁兹派的领导人——十分有趣的经历。

——约翰·麦凯恩，美国议员

和几个朋友出去玩——“砰”马提尼——准备好看圣诞特辑。东部10点，中部9点。然后一起欢唱圣诞颂歌！

——奥普拉·温弗瑞，媒体大亨和电视节目主持人

是时候喝瓶红酒然后想想下一场巡演了。圣路易斯棒极了。我的眼线蹭到了膝盖上，胳膊肘上有血。不高兴

——Lady Gaga，歌手和流行歌曲作家

大约10万年以前，我们的祖先开始说话。大约5000年前，人类开始写字。在过去的150年里，我们的接收方式从电报、电台、电视转变为电子邮件、短信、博客和其他社交媒体。虽然沟通方式发生了改变，但我们还是不得不和周围的人交流自己的想法、经验和情感。

从2006年起，我们开始使用推特网。任何一个有推特账户的人都可以发简短的消息，也称“推文”，发出的推文可以很快地几乎被所有人看到。分分钟内，你就可以知道你的朋友，或是世界有名的明星们在想些什么。很多读者也许会好奇，为什么大家会想要这么做。然而，如果一旦让自己沉浸在推特的世界中，你也会发现它吸引人的地方。

我们回过头来看这一章开头的那四条推文。从某种程度上来说，这几条推文与日常沟通没有什么不同。我们在餐厅吃饭时，偶然听到的邻桌谈话内容可能就与此相似。这些不同的人在向他人表达什么？帕丽斯·希尔顿只是在大声地打招呼。约翰·麦凯恩描述了他和一位黎巴嫩重要领导人的会面。奥普拉·温弗瑞告诉了我们她晚上的计划。Lady Gaga想让我们知道她已经进入下一次巡演的状态了。

但是，这些推文包含的内容要多于它们的作者想要表达的意思。每一条都像一个指纹。比如说，如果这是一个多项选择测试题，题的要求是将这些推文和作者进行匹配，那大部分人都可以拿满分。即使你从未听说过这几个人，仅凭“媒体人物”“美国议员”等这些标签也能推测出哪条推文对应哪个作者。

通过这些推文还能深入了解每个作者的想法和个性。感叹号和表情符号表达了希尔顿持续高昂的情绪。麦凯恩用一些大词和他的世故给读者留下深刻印象。温弗瑞像一个完美的售货员，“无意地”说出了圣诞特辑（实际上是她主持节目的特辑）直播的时间。Lady Gaga想表达的是她有一点狂野，但是也很体贴。从使用代词的方式来看，她有点忧郁倾向。

如果我们开始分析他们每个人更多的推文，就会对他们的动机、恐惧、情感和他们与自己及别人的沟通方式有更加丰富的了解。每个人使用语言的方式都是独特的。有些人，像是Lady Gaga，他们的沟通方式更倾向于表达个人——从他们使用的词汇，如“我”，可以看出他们是自我反映型的。其他人，像是约翰·麦凯恩，暴露了他在与人沟通方面有很大的问题。

如果你想要尝试一下基于推特文稿的快速性格分析工具，可以登录一下我和同事共同创建的试验性网站：www.analyzewords.com 。

通常情况下，那些最能透露信息的词是最短的和最容易被人遗忘的词。代词（如你、我、我们、他们 ）、冠词（某个、这个 ）、介词（如至、为、多于 ）和其他隐性词汇能透露出我们是谁。这就是本书将要讲述的内容。

从我们的祖先说出第一个句子到帕丽斯·希尔顿发推文打招呼经历了漫长的过程。多亏当前的技术革命，让我们拥有了工具来分析推文、脸书、邮件、书信、书籍和日常生活使用的文字。有史以来，这是第一次我们能够用计算机来确定日常语言是如何反映出一个人的社会地位和心理状态的。

比如说，谁会预想到，如果一个高中生在大学入学申请中使用的动词很多，那他有可能入学后成绩比较差？或者一个诗人在他的诗中过多地使用“我”这个词，那他自杀的可能性会比较高？又或者可以通过某个领导人所使用的代词来推断他是否会带领他的国家陷入战争？通过仔细观察人们传达思想的方式，我们可以开始了解到人们的性格、情感和社会关系。

当社会心理学和语言学相遇

在描述词汇的秘密之前，也许有必要说一下本书的作者。那就是我。我是一个社会心理学家，我对语言的兴趣来源于偶然。正如你将要看到的，本书真正的重点是人而不是语言本身。当然，词汇和语言是引人入胜的话题。在一个社会心理学家的眼中，语言甚至更加吸引人，因为它为通往人的内心深处提供了线索。

与此相关的背景是，我早期的事业都涉及健康、情感和创伤经历。20世纪80年代初期，一个偶然的发现深深吸引了我。那些报告了自己经历过严重创伤的人和把创伤经历作为秘密的人的健康问题比那些公开讨论自己精神创伤的人要严重得多。谁会保守一个对自己有害的秘密？更重要的是，如果让一个人揭开自己心中的情感秘密，他的健康状况会有所提高吗？我和我的学生很快发现，答案是肯定的。

我们开始做实验，实验中我们要求被试者连续三天或四天每天花15～20分钟的时间来记述自己的创伤经历。与每天记述非情感类话题的被试者相比，这些记述创伤经历的被试者的身体健康有所改善。后续研究发现，情感写作能够提高免疫功能、降低血压、减少压抑感和提高情绪。现在，在距离第一次写作实验25年后，全世界已经有两百多个类似的写作研究在进行。虽然效果不是特别明显，但是将情感波动转化为文字的行为确实有助于改善身心健康。

◎如何解释文字的力量

为什么写作会起作用？一些科学家认为，重复性地面对自己的痛苦情绪，最终会降低它们带来的影响，因为我们接受了它们。另一些科学家指出，对痛苦的沉思和未完成情结会带来不健康的影响。很多受到精神创伤的人会不停地在脑海中重复发生过的事情，妄图弄明白所受的痛苦。无止境地思考他们的情绪波动会扰乱他们的睡眠，并使其无法专注于自己的工作和人际关系。根据这个观点，把创伤经历写出来，可以使人们理解曾经的经历并从中找到意义，还可以帮助他们解决情感上的波动。

答案并不简单。我相信，当人们写下他们的痛苦经历时，一些健康的变化正在发生，包括思维方式、情绪反应、大脑活动、睡眠和健康行为的改变等。即使发现了写作对一个人起作用的原因，也不一定能够解释为什么写作会对其他人起作用。

之前的写作研究者没有考虑到的是，人们用语言来描述他们的个人情感剧变，也许表达性写作的关键隐藏在人们的文字背后，隐藏在他们真正

要表达的内容中。人们所写的故事是有力量的，常常让人难以忘怀。在几乎所有的项目中，参与者会记述身体虐待、性虐待、情感上的虐待、离婚问题、毒品问题、酒精问题、自杀和可怕的意外，以及失败、羞辱和痛苦的感觉。不仅有各种各样有力量的故事，而且每个人的表达方式都非常不同。有些人幽默，有些人充满愤怒，还有些人冰冷、超然且不带感情。

如果一组临床心理学家或只是普通的人阅读这些文章，他们可以解读出什么样的文章预示着身体健康的改善吗？我们尝试了这样做，得到的答案是否定的。这些故事都太复杂，即使是最认真的读者也无法确定这些令人心碎的故事中哪些部分是最有意义的。我们需要其他方法来揭秘表达性写作有效性背后的原因。

◎计算机革命和LIWC的诞生

那是1991年，计算机技术革命正在顺利进行。20世纪六七十年代，普林斯顿大学、哈佛大学和麻省理工学院在语言研究的计算机分析方面取得了一些重大突破。无疑，有了这种新技术，我可以得到一个能够分析情感文章的计算机程序。不需要判断，不需要费脑细胞，我可以一键获得想要的一些答案。

不幸的是，那个时候没有能供我使用的计算机程序。“写一个程序能有多难呢？”我问我自己。巧的是，有一个一直以来都是专业程序员的研究生新生刚刚加入了我的研究团队。“玛莎”，偶然间我跟她讲，“我有一个好主意——编写一个新程序，这个过程只需要大约三周的时间。”玛莎·弗朗西斯是一个在社会心理学方面有天赋且富有创造力的程序员，因为她完全不知道这是用来做什么的，而且程序很快就写完了。就这样，“三周项目”有了自己的生命。三年后，我们终于推出了这一计算机程序的第一版，我们称之为“语言获得和词汇计数”（Linguistic Inquiry and Word Count），或LIWC（发音同“卢克”）。

LIWC背后的想法是，无论是创伤写作或是日常交流，人们的语言可以反映人们的感受，我们通过简单的词汇计数可以深入了解他们的情感状态。我们猜想，愤怒的人会使用与愤怒相关的词汇，悲伤的人会使用与悲伤相关的词汇。在记述创伤经历时，被试者的情感状态会通过他们使用的与情感相关的词汇反映出来。

所以，在开发LIWC程序时，我们创造了一系列的词典，旨在捕捉到不同的心理状态。例如，我们创建了愤怒词典，现在已经包含180多个与愤怒相关的词，如仇恨、暴怒、杀害、鞭打、复仇 等。词典中还包含了一些词干，如杀 ，这样所有包含杀字的词都会被计数在内，如杀人犯、杀害、谋杀、被杀 。我们之后还创建了与悲伤、焦虑、积极情绪和其他情绪

状态相关的词典。

除了情感基调的不同以外，被试者所写的创伤文章还在很多方面呈现出不同。为了包含的更全面，我们还创建了可以统计其他类型词汇的词典，如不同类型代词的使用（如第一人称代词单数：我、我的 ）、冠词（某个、这个 ）、与因果思维相关的词汇（引起、因为、原因、原理 ）等。不知不觉，我们已经创建了将近80个不同的词典，我们认为这些词典几乎囊括了人们日常中会使用到的所有词汇。

之所以LIWC程序的完成花费了我们三年的时间，就是因为创建词典的过程是漫长的。我们雇了一群学生来评估词典中的每一个词。举个例子，“沮丧”这个词是否应当被列入愤怒词典中呢？只有学生评委小组全数通过才可以被列入其中（此例的结果是全数通过）。

多亏了玛莎的编程技术和学生评委花费的成千上万个小时，才使得LIWC终于可以运行使用。这款程序最终可以在瞬间完成对电子文档的分析，并按照每一个词典来统计词频数。LIWC最新版本可以在数秒内完成上千个电子文件的分析。虽然我们最初的研究重点是创伤文章，但是后来我们将研究范围扩展到了诗歌、小说、博客、推文、书信、短信、对话记录和其他任何包含文字的文档。

为了介绍这一计数程序是如何运行的，我们先来看一下刘易斯·卡罗尔的《爱丽丝梦游仙境》（Alice’Adventures in Wonderland）的前两句：

爱丽丝开始觉得，和姐姐一起坐在岸边什么都不做非常的无聊，她偷偷看了一两眼姐姐正在读的书，但是里面没有任何的图片或对话，“这本书有什么好看的？”爱丽丝想，“连图片和对话都没有。”

所以她想（尽她所能，因为天气热的缘故，她感到非常的困和迷糊），做一个菊花花环获得的快乐是否值得我站起来去采菊花呢，这时一只粉红眼睛的白兔突然跑了过来。

首先，LIWC会先统计文章的总词数，这段的词数是162。[1] 然后它会单独查看每一个词，决定它应该被归在哪一类词典中。所以，比如说，LIWC会先查看“爱丽丝”，但是在任何一个词典中都找不到这个词。但是如果一个词出现在了多个词典中，如动词词典、助动词词典和过去时态动词词典中都有它，那么每个词典中的计数都会加1。LIWC会继续它的工作，接下来查看下一个词，“开始”，定位为时间词典，然后再查看下一个词。最后，当评估完文章中的所有词，并将它们都指派给相关词典后，LIWC会计算每一类词占总数的百分比。

在分析文本方面，LIWC比专家拥有更多的优势。LIWC程序是百分百可信的，因为如果多次分析同一文档，每一次的结果都是相同的。它的速度非常快，能够在20秒内完成对莎士比亚作品集的分析。而且，一个人的文章的分析结果可以直接与其他人的做对比。尽管计数程序有这些令人钦佩的特点，但它也非常笨。它无法识别出讽刺和挖苦，也非常缺乏幽默感。最糟糕的是，它无法捕捉到语言环境。比如说，一个词根据它的用法不同，可以有完全不同的意思。

想想“疯”这个词。LIWC程序将“疯”归类到愤怒和消极情绪的词典中。如果一个人说：“你疯了吗，竟然亲吻我的男朋友。”那么LIWC把“疯”归入愤怒词典中是没有问题的。但如果一个人说：“我疯狂地爱上了我的男朋友。”那么LIWC还把“疯”归入愤怒词典中就是错误的。在这个例子中，“疯”不是表示愤怒，而是表示“疯狂得高兴”。或者，我们回到《爱丽丝梦游仙境》，“疯茶派对”也不意味着是有敌意的。相反，“疯”在这个语境中的意思是“奇特”或者“反常”。

与其他计数系统一样，LIWC也会出错。它是一个概率系统。有时候它的分类是正确的，有时候不是正确的。我们已经进行了足够多的实验来确定，通常它的统计结果是正确的，而且好消息是，它分析的词越多，结果会越准确。另一个好消息是，研究人员正在开发更智能的词计数程序，最终会将句法、语法和上下文考虑在内。

考虑到当前的艺术状态，难怪那些严肃的语言学家和文学学者对词计数程序会有些反感。语言学家关心语言，文学学者关心文学。而我，以我自己的方式关心着这两点。但我真正感兴趣的是文字的使用和人们的心理状态之间的关系。我们能否确定出哪些语言特征能够揭示人们思考方式呢？如果可以，我们能否通过这一信息以有益的方式来改变人们的思维呢？

[1] 原文统计的词数为英语单词数，共113，此处162是译后的中文词总数。——译者注

通过计算机程序识别健康的写作

开发LIWC是为了弄明白，通过人们记述创伤经历的方式是否可以预测到后续健康的改善情况。换一种说法就是，我们能否通过LIWC来识别健康的写作。

在LIWC的开发过程中，已经有一些表达性写作的研究可以解答这个问题。在我的实验室里，有三项实验的被试是大学生。其他实验是由我的同事组织进行的，被试的类型更加多样化，包括最高安全级别的犯人、一组新西兰学医的学生和一群最近刚失业的高级工程师。现在我们有了六组不同的文章样本，来自不同的人，可以使用我们的计算机程序来寻找哪些类别的词是与健康写作相关的。三个重要的结果出现了。

◎积极情绪的重要性

通过对创伤文章中表示积极情绪的词（如爱、关心、高兴 ）和表示消极情绪的词（如悲伤、痛苦、愤怒 ）的计数情况，可以大致推断出一个人的情感状态。六组文章的研究结果令人感到意外。总的来说，在描述自己的情感波动时使用越多表示积极情绪词的人，在实验进行数周或数月后，健康状况好转越多。

消极情绪呈现出完全不同的模式。健康状况好转最明显的人使用消极情绪词汇的比例适中。也就是说，使用消极情绪词汇非常多的人没有从中受益——他们好像被不幸所淹没。同样地，那些使用消极情绪词汇非常少的人也没有从中受益，这也许说明了他们在否认情感对他们产生的影响。情感的分析结果显示，从创伤写作中获益的最佳方式是承认消极面，并赞美积极面。

◎构建故事的重要性

LIWC呈现的令人激动的一方面是，它能够识别反映人们积极思考程度的词类。两组认知维度的词汇，一组是洞察力和自我反省词汇（如思考、认识到、相信 ），另一组是因果关系词汇（如因为、影响、原理 ）。健康状况好转最明显的人最初使用认知类词汇的比例非常低，但在实验进行的四天中，使用量逐渐增加。认知词汇所占比例不是最重要的，重要的是比例从第一天到最后一天有所增长。在某种程度上，使用洞察力和因果关系词汇对人们构建一个连贯的创伤经历是必需的。在写作的第一阶段，人们通常记述得非常混乱。然而，随着日复一日的记述，他们开始理解自己所写的文字。他们能够更好地理解文字，部分体现为他们使用认知词汇的方式。

这些结果表示，拥有一个解释痛苦经历的连贯故事不一定像构建一个连贯故事一样有用。这有助于解释一个我的个人观察，这一观察困扰了我多年。第一次写作研究发表后，我的文章常被媒体作为专题来报道。在鸡尾酒会或非正式聚会上，我发现有时我像创伤磁铁。那些知道我的研究的人会主动靠近，来给我讲述他们可怕的人生经历。其中有很多人的身体状况也很差。起初我认为，谈论他们的故事会对他们有益处。然而，我会在数月后的另外一个聚会中看到同一个人，给我讲述一模一样的故事，并且他们的健康状况没有什么变化。

词汇计数研究揭示了这一问题。那些讲述他们创伤故事的人基本上是在不断重复同样的故事。故事没有变，对故事的理解也没有加深。以同样的方式重复讲述同样的故事就像反复式思考，这是典型的抑郁症状。

这里给大家上重要的一课。如果你被生活中的情感波动所困扰，尝试写出来或向他人倾诉。但是，如果你发现自己为了摆脱悲痛而不断重复讲述同样的故事时，请重新考虑你的方式。尝试使用完全不同的方式来记述或讲述你的创伤经历。一个更加超然的讲述者会如何描述所发生的事情？还会有什么其他的方式来解释这件事呢？研究表明，如果你成功了，你会睡得更好，身体健康状况会有所改善，你会发现自己比以前更快乐，更容易克服情感波动。

◎改变视角的重要性

多亏LIWC程序，我们发现情感写作的三个方面预示着人们身心健康的改善：强调情感波动中的积极部分；承认消极部分；在写作的日子中构建一个故事。不久，更多复杂的分析揭示了前人从未看到的语言使用的另一面。

20世纪90年代，一组科罗拉多大学的研究者推出了名为“潜在语义分析”（Latent Semantic Analysis，LSA）的计算机程序，它可以在文章内或不同的文章间寻找词汇的使用模式。LSA的妙处在于，它可以从数学角度比较任何两个文本间的相似程度。那么从表面上看，我们可以确定，每天记述同样话题的人是否会比每天记述不同话题的人从中获益更多。这是一个宏伟的想法，但任凭我们努力，也找不到充分的证据来支持它。

我的研究生学生夏洛克·坎贝尔曾花了差不多一年的时间来让LSA程序运行。我和他越考虑LSA项目，越觉得我们思考语言的方式是错误的。与其分析人们写了什么内容，为什么不分析他们的语言风格呢？为了做到这点，我们需要完全改变LSA程序。我们不再通过名词、常见动词和形容词来分析文章的内容，取而代之，我们让程序关注那些可以体现写作风格的词。我们所了解到的是，写作风格通常是通过功能词，包括代词、介词、

冠词，以及少量的较短但是比较常用的词汇来体现的。

结果是惊人的。（好吧，如果你不是计算机语言学家，那“惊人”这个词可能有一点夸张，你需要在现场才能体会到。）在写作过程中，一个人使用功能词的方式转变得越多，他的健康情况改善得越多。当我们开始关注不同类别的功能词时，一组特殊的元素脱颖而出，占据了非常重要的地位，它就是人称代词。更具体地讲，与其他代词相比（如我们、你、她、他们），人们使用第一人称单数的改变越明显（如我、我的），他们的健康状况改善得越明显。人称代词产生的影响很大，关于它的研究也在不断进行。

在这个计算机程序上花费了一年多的时间后，夏洛克感到非常激动。他指出，我们发现了“代词的奥秘”，他为此感到非常自豪。

这一发现可能听起来深奥，但在现实生活中并非如此。那些健康状况有所好转的人在某篇文章中使用“我”字的频率很高，下一篇文章中使用其他代词的频率很高，然后在接下来的文章中使用频率反复切换。换句话说，健康的人在描述完自己的想法和感受后，会去探究一下别人身上发生的事情，然后再回到自己身上。

这种视角的转换在心理治疗中非常常见。如果一个男性见到他的心理治疗师以后，一直在抱怨他妻子的行为，她说了什么，如何冷漠等，心理治疗师很可能会在几分钟后打断他：“你一直在谈论你的妻子，却没有说说你自己。这些事情发生时你的感受是什么样的？”同样地，如果另一个客户（比如是事件中的女主角）在和她的心理治疗师谈论婚姻方面的问题时，花费了很长的时间谈论她自己的想法、感受和行为，却从来不谈论她的伴侣，那心理治疗师也可能会以同样的方式转移话题：“你说了很多自己的感受，你觉得你丈夫当时是怎么想的？”也许就像好的疗法，健康的写作可以让人从不同的视角来看问题。

语言像镜子，也是工具

停留一分钟，然后想一下我们发现的关于表达性写作的结论。记述情感波动可以改善一个人的身心健康，然而不是所有人都能从这一练习中获益。那些健康得到改善的人与没有改善的人的表达方式有所不同。健康的写作涉及使用积极情绪词汇，适度使用负面情绪词汇，认知词汇不断增多，以及代词使用的变化。这些表现在每天的写作中是这样的：会从写作中获益的人的表达更积极向上，会承认事情的消极面，为自己的经历构建一个有意义的故事，并且可以在写作时使用不同的视角。

然而，最让人吃惊的是，这些结论是从几乎不被人注意的隐性词汇的使用中反映出来的。隐性词汇一直在那里，反映着人们思考方式发生的关键性变化。

这些关于语言的发现确实有趣，但我们可以很好地利用它们吗？如果我们把人带进实验室，鼓励他们写作时使用积极情绪词汇，多使用认知词汇，改变他们使用人称代词的方式，他们的健康会改善吗？换句话说，语言反映一个人心理状态的改变，还是会导致其心理状态的改变？

多年来，有一些研究在尝试回答这一问题。谢丽尔·休斯是之前我在南卫理公会大学的学生，在一项精心设计的实验中，她给被试的学生提供了一个词汇列表，要求他们在表达性写作中使用这些词汇。有些收到的是积极情绪词汇表，有些是消极情绪词汇表；有些收到了认知词汇表，另一些没有认知词汇表。她成功地操纵了被试者按着她预期的方向使用词汇，结果却没有对健康产生任何影响。还有一些其他巧妙的尝试，以改变人们在与情感话题相关的写作中使用认知词汇的频率，或使用人称代词的类别。目前的结果具有一定的信服力：词汇使用通常反映人们心理状态的改变，而不是影响或使心理状态发生改变。

词汇使用可以反映我们的想法和感受，这不是一个令人吃惊的发现。但是这些结果表明，我们是可以通过词汇分析来改变人们思维方式的。回想一下，健康写作的特征在于某类词汇的增加，比如“因为、引起、影响、原因”，以及与认知相关的词汇。简单地要求人们在写作时更多地使用这些词汇是没有意义的，因为写作者只是关注了词汇，而不是关注它们的根本目的。但是，如果我们鼓励人们记述自己的创伤经历，并构建一个有意义的故事，那么他们的写作基调会更有活力。他们开始以局外人的视角来看待自己的经历。他们在写作时投入的认知工作使得故事写得更好，这更有可能帮助他们走出创伤。

词汇分析告诉我们人们如何思考的同时，给我们提供了一种指导人们思考的方式。语言可以是镜子，也可以是工具。

***

有趣的是，到目前为止，关于表达性写作的发现都只是使用LIWC的众多方法的补充说明。当玛莎和我开发这款软件时，我们的目标很简单：通过查看词汇来识别健康写作。但是，我们的时机非常有利。在没有计算机技术和可用大型文本数据库的情况下，语言学家也许已经研究过词汇，心理学家和内科医生也许已经研究过健康，但是还没有人将两者自然地结合起来。

现在，任何人都可以在很短的时间内访问上千人使用的数以百万计的词汇。作为一个爱摆弄数字和统计学的人，我在LIWC里发现了游乐场。我经常发现自己在深夜分析文本，只是为了看看谁在使用什么样的语言。比如说，大部分表达性写作的研究包括男女两性，那么男女在语言的使用方式上是否存在差异呢？是的，存在差异，而且差异很明显，但是这并没有什么意义。所以我做了任何尝试在深夜解决问题的人都会做的事情：我忽略了它。然而，我在第二天晚上又上传了另外一批文档，发现了同样奇特的差异。我不会现在告诉你这些发现，因为那会毁掉第3章给你带来的兴奋感。但是请相信我，这些差异非常明显且出乎意料，而且如果你仔细思考，就会发现它们是完全合理的。

随着我上传的数据越来越多，重复的词汇模式不断出现。词汇使用差异不仅存在于男女之间，而且不同年龄的人之间也存在很大差异。不同社会地位的人之间使用词汇的方式也存在很大差异。情感状态、诚实程度、性格类型、语言正式度、领导能力、关系的好坏等方面的不同，都会影响语言的使用方式。语言使用几乎与我研究过的每一个社会心理学维度相关。尤其耐人寻味的是，大多数的文字差异是由英文中那些最普通、最易被人忽略的词引起的。

如果在30年前有人告诉我说，发现代词的奥秘将会成为你事业的高峰，那我一定已经转行了。但是现在我相信，通过理解人们何时使用、如何使用功能词，我们可以了解影响人们行为的社会状况和心理状况，包括他们和朋友、家庭成员的关系，以及我们和商务人员或外界有效沟通的能力。

如果一个朋友的语言突然发生了变化，我们也许可以确定他是否正在感到失落、生气或者被欺骗。如果你是一个档案保管员或调查员，那会有助于你通过解码功能词来识别文字的真实作者或文字背后的动机。

然而，对我来说最有趣的是，分析说话或写作时使用的词汇对于认识我们自己非常有帮助。在接下来的章节中你会看到很多例子，我分析了自己的电子邮件、推荐信和与家人朋友的日常交流。通过仔细观察文字，我发现了改善与朋友和家人关系的方式，那就是做一个更好的老师，成为一个更好的领导者。我没有从根本上改变自己说话或写作方式，这些文字分析指出了我天然存在的一些缺点，我在努力去改掉这些缺点。

最后，本书是我和我的学生、同事研究人们使用语言方式的旅行指南。一路走来，我和各类心理学研究者一起工作，还有社会学家、语言学家、计算机科学家、沟通专家、律师、历史学家、FBI特工以及其他调查员……你知道的。指导大部分研究的价值取向是好奇心理和游戏性。如果你是一个严肃的语言学家，本书可能会让你失望，甚至会惹怒你。如果你是热爱文字本身，那我可能没办法共享你的这份崇敬。（事实上，我有一篇论文讲到自杀型诗人使用代词的方式和非自杀型诗人不一样，这篇论文发表之后，有一个酒醉的诗人在我家的派对上拿着黄油刀来威胁我。）从根本上来讲，我感兴趣的是心理学和社会行为。在我的世界里，词汇是通往人类内心的窗户，它精彩而发人深省地揭开了语言和这个世界的联系。

第2章　重要的不是内容，而是语言风格

使用冠词频率高的人更有组织性，情感更稳定，更加谨慎，政治上更加保守，年纪上更大一些。

不是所有的词功能都是一样的。在任何给定的句子中，有些词提供了基本的内容和意义，其他的词只是辅助性词汇。然而讽刺的是，这些默默无闻的词比那些有含义的词更能体现一个人性格特点。本书的一个中心思想就是，说话内容应该与说话风格区分开来。更进一步，词汇可以反映语言风格，而语言风格可以揭示一个人的性格、社会关系和心理状态。

本章介绍的词汇分析的整体逻辑，是本书下面所有内容的基础。如果你只是想了解不同的词汇类别如何反映出一个人的性格、伪装方式和其心理状态，那么你可以跳过此章节。你可能会感到遗憾，但没关系，反正你永远也不会发现。

***

我们以一个简单的练习开始，这有助于你对不同的语言是如何传达信息的有一个初步的概念。仔细看一下这幅图。这两个人是谁？他们在干什么？他们有什么想法和感受，他们在关注什么？



你会如何描写或如何谈论这幅图？停顿一下，然后自己描述一下这幅图。你甚至可以在纸上记下自己是如何描述的，以便在阅读这章的过程中回头来看你写的内容。

事实上，上千人已经描述过这幅图，并成为各种心理实验的一部分。人们讲述的故事千差万别。有些人认为这是两个女性，有些认为是一男一

女或者两个男性。有些人将这两个人编入了关于善与恶、智慧与青春的故事中，或只是不同两代人之间的家庭故事中。

尽管忽略故事和主题的不同，但人们描写故事的方式更加惊人。举个例子，下面是三个大学生所描写的故事的第一句话，我们来读一下：

学生1：在上述画面中，一个老年妇女正要和一个看上去高人一等并精于算计的中年女子说话。

学生2：我看到一个老妇人在回想她当年年轻美丽时候的样子。

学生3：那个老妇人是个巫婆之类的人物。她看上去正在哄骗这个年轻人去做件什么事。

现在，我们来仔细地读一下每个人的描述。你能大概对每一个人有个初步的印象吗？如果让你和其中一个人喝咖啡，你会选谁呢？你最信任哪一个人？哪些因素影响了你的答案？虽然这三位同学看到的是同一幅图，但是他们的理解各不相同。然而更重要的是，他们是如何用语言表达自己感受的。即使是在这么简单的语句中，你也能对这几个学生的性格有个大概的印象。第一个人呆板冷漠，会不自觉地使用一些长难词。第二个人更加个人化、更加温情。第三个人比另两个人更加随意，似乎没有认真对待这个测试。

这三个人都将他们个性的一部分融合在了自己的写作风格中。他们使用词汇几乎是下意识的，通过这些词汇，我们可以了解他们是谁，可以了解他们对自己和他人的看法。我们可以做一个合理的推测，学生1是个男孩，另外两个是女孩。学生1的平均成绩可能高于其他两人，尽管他的社交生活可能比较痛苦。学生2是三个人中最有可能情绪低落的。学生3可能在学校的表现不是很好——也许花太多时间和朋友厮混，还常常喝很多酒。

这些不是有根据的推测，而是有根据的判断。这个根据是，人们日常生活所使用的语言可以在很大程度上表明他们的个性、年龄、性别、社会阶层、压力水平、生理活动和社会关系。语言为了解一个人提供了线索。通过分析这些线索，我们看到了每个作者的内心世界。

基于你对这幅图的描述，大家能判断出什么呢？我们是如何知道学生1是一个社交孤立的男生，而学生2是一个沮丧的女生呢？秘密就在于如何区分语言内容和语言风格间的不同。回过头来看上面的三个叙述，写作的内容确实不同，但是更加不同的是他们表达的方式。这就存在一个有意义的差别，那就是语言内容和语言风格的差别。

用词风格是心理的钥匙

什么是风格？现代人格心理学的创始人高尔顿·奥尔波特提出了这个问题，尝试定义人类的本质区别。他指出，人们做的几乎每一件事都可以展示自己。有些人走路很快，不摆动手臂；有些像是脚下踩了弹力球一样跳着走；还有些人从容漫步，歪斜着走或是沉重地走。高尔顿认为，走路风格的不同是体现人们不同的一种方式。然而，他们在穿衣、饮食和剥橙子上也不同。风格也许不会告诉我们一个人在哪行走，他有多饿，或者他喜欢的水果是什么，但它是通往人们性格、态度和社会关系的窗户。

语言风格也不例外。人们的说话和写作方式能够显示一个人的性格。难点在于如何确定哪些因素决定语言风格。有趣的是，语言学家、高中英语老师和大自然为我们提供了反应语言风格和语言内容的提示。

实义词 是指用来标记某个物品或动作且具有文化共享含义的词汇。就我们的研究目的而言，实义词包含以下几种。

名词　如桌子、叔叔、正义、怪币

规则动词和行为动词　如爱、走、躲

大多数修饰语　如形容词（蓝色的、快的、令人垂涎的）和副词（悲伤地、饥饿地）

实义词对于向某人传递一个概念来说是非常必要的。考虑一下下面三个人关于图片进行的简短描述。

学生1：在上述画面中，一个年老的妇女正要和一个看上去高人一等并精于算计的中年女子说话。

学生2：我看到一个老妇人在回想她当年年轻美丽时候的样子。

学生3：那个老妇人是个巫婆之类的人物。她看上去正在哄骗这个年轻人去做件什么事。

想象一下你正在和一个英语很烂的人交谈。那个人正在试图描述这幅图。你能够理解的就只有突出显示的这些实义词。事实是，你可以通过这些与内容相关的词汇很好地了解到说话人想要表达的事情。这很好。实义词应该传递含义。不可否认，这不会是一场非常令人满意的互动，但是你大概可以确定说话人脑袋里想的是什么。

风格词（或功能词）是连接和组织实义词的词汇。虽然这个定义有点狡猾，但是大部分的风格词归属于一大类，它们被称为功能词、隐性词甚至是垃圾词汇。一个好的理解功能词的方式是：它们自身是没有任何含义的。例如，实义词“桌子”可以触发人们脑海中的想象，同样，“散步”“蓝色”和“小虫”也可以触发人们的想象。现在试着想象一下“那个”“因为”或者“确实”，甚至是“你”。我们可能会在大多数的句子中使用这样的词汇，但它们本身几乎是无用的。

多数情况下，功能词包括：

类别	示例
代词	我、她、它（I,she,it）
冠词	某个、这个（a,an,the）
介词	之上、和、在……里面、为（up,with,in,for）
助动词	是、未、已经（is,don't,have）
否定词	不、不是、从不（no,not,never）
连词	但是、而且、因为（but,and,because）
数量词	少量、一些、多数（few,some,most）
常用副词	非常、确实（very,really）

为了更好地领会到这些常被误会的词汇的重要性，我们来回过头看一下这三个人是怎么描述图片的。

学生1：在上述画面中，一个老年妇女正要和一个看上去高人一等并精于算计的中年女子说话。

学生2：我看到一个老妇人在回想她当年年轻美丽时候的样子。

学生3：那个老妇人是个巫婆之类的人物。她看上去正在哄骗这个年轻人去做件什么事。

现在想象一下一个人尝试向你描述这幅图片时只会使用突出的这些功能词，那你肯定不知道这个人想要说什么。

那为何功能词还那么重要呢？因为代词、介词和其他功能词是心理的

钥匙。好吧，也许这样说有点言过其实，但是请听我继续讲。隐性词汇具有以下特点：

·使用频率很高

·短且难以被察觉

·在大脑中的处理过程与实义词完全不同

·非常非常社会化

每项特征都有助于解释为什么功能词是重要的心理关口，与此同时解释了为什么很少有人仔细地研究它们。另外，隐性词汇的确对风格的影响非常大。是时候让这些人们脱口而出并总被遗忘的小词得到应有的关注了。

◎日常生活中的功能词：它们无处不在

1863年，毁灭性的葛底斯堡战役结束四个月后，亚伯拉罕·林肯发表了美国史上最著名的演讲之一。

俯瞰战死了7500位战士的战场，林肯简短的发言重新塑造了美国内战。我们来快速地阅读一下他的演讲，以便对他所说的内容有个印象。

87年前，我们的先辈们在这个大陆上创立了一个新国家，它孕育于自由之中，奉行一切人生来平等的原则。

现在我们正从事一场伟大的内战，以考验这个国家，或者说以考验任何一个孕育于自由而奉行上述原则的国家是否能够长久存在下去。我们在这场战争中的一个伟大战场上集合。烈士们为使这个国家能够生存下去而献出了自己的生命，我们在此集会是为了把这个战场的一部分奉献给他们作为最后安息之所。我们这样做是完全应该而且非常恰当的。

但是，从更广泛的意义上来说，这块土地我们不能够奉献，我们不能够圣化，我们不能够神化。曾在这里战斗过的勇士们，活着的和去世的，已经把这块土地神圣化了，这远不是我们微薄的力量所能增减的。全世界将很少注意到，也不会长久地记起我们今天在这里所说的话，但全世界永远不会忘记勇士们在这里所做过的事。

毋宁说，倒是我们这些还活着的人，应该在这里把自己奉献于勇士们已经如此崇高地向前推进但尚未完成的事业。倒是我们应该在这里把自己奉献于仍然留在我们面前的伟大任务，以便使我们从这些光荣的死者身上

汲取更多的献身精神，来完成他们已经完全彻底为之献身的事业；以便使我们在这里下定最大的决心，不让这些死者白白牺牲；以便使国家在上帝福佑下得到自由的新生，并且使这个民有、民治、民享的政府永世长存。[1]

现在，闭上眼睛回想一下这篇演讲的内容。哪个词出现的频率最高？在脑海中尝试回忆一下，在林肯创作的这样一篇有力的演讲中哪个词的使用频率是最高的。我是认真的。闭上你的眼睛，然后在脑海中列出演讲中使用频率最高的词汇。

好，现在你可以睁开眼睛了。被要求做这件事情的大多数人是毫无戒心的，他们想到的词多是“民族”“战争”“人类”，或者可能是“死者”。当你得知使用频率最高的是功能词而非实义词的时候，可能不会感到太吃惊。在这篇特定的演讲中，使用频率最高的词是“的”。其他使用较频繁的词汇还有：这里、一个、为、能够、我们、了、已经、不。事实上，在林肯这篇精心创作的演讲中，使用频率较高的前几名都是这些小词。这样一篇伟大的演讲中，小的不起眼的词汇占据了很大的比例，这是一件值得注意的事情。

隐性词汇虽然很少，却占据了我们所听、所读和所说词汇的很大一部分。在过去的20多年里，我和我的同事们收集了大型的文本文件，包含成千上万的日常对话、书籍、互联网博客、音乐歌词、维基百科条目等，包含数十亿的文字。尽管根据人们说话或写作的内容不同，词汇的使用会存在一些变化，但是你会惊奇地发现，功能词在各种类型的文本中都很常用。

花一分钟仔细看一下下面的表格。这个列表是基于大的语言库统计出的英语中最常使用的20个词。统计横跨书面语和口语，例如，词语“我”（I）占据了我们使用的所有词汇的3.6%。如果将这20个词都考虑在内，那么它们总共占据了人们所用、所读和所听的所有词汇的近30%。

请注意，表2-1中的所有词都很短，而且由代词、介词、连词、冠词和助动词组成。如果我们将列表扩展为所有常用的英文隐性词汇或者说英文功能词，那么这个列表大约会包含450个词。事实上，这450个词占据我们使用的所有英文词汇的一半还多（55%）。

表2-1　口语和书面语中最常用的单词列表

排名	词	占比 (%)
1	我（I）	3.64
2	这（the）	3.48
3	和（and）	2.92
4	至（to）	2.91
5	某个（a）	1.94
6	属于（of）	1.83
7	那个（that）	1.48
8	在……里面（in）	1.29
9	它（it）	1.19
10	我的（my）	1.08
11	是（is）	1.06
12	你（you）	1.05
13	是（was）	1.01
14	为（for）	0.80
15	已经（have）	0.70
16	和……一起（with）	0.67
17	他（he）	0.66
18	我（me）	0.64
19	在……上面（on）	0.63
20	但是（but）	0.62

换一个角度来看这个问题，普通说英语的人所具备的词汇量令人钦佩，大概是10万个词。这意味着只有非常少的一部分词是与语言风格相关的——大约只占所有词汇的0.04%。而剩下的99.96%都是由实义词构成。其他语言中，这个比例也是同等水平的，例如德语、西班牙语、土耳其语、阿拉伯语、韩语和其他我们已经研究过的语言。在所有语言中，功能词都只占很小的比例，但是与实义词相比，功能词的使用比例惊人的高，而实义词使用的比例却非常低。

稍微思索一下这些数字意味着什么。如果你打算学习一门新的语言，比如德语或芬兰语，你可以只花大概一下午的时间就能够掌握一大半的德语或芬兰语。几乎任何一个人都能够以最少的培训掌握前100个隐性词汇。到了傍晚时分，你就可以坐下来阅读德语报纸或芬兰语的哲学文本了，然后发现你认识其中一半的词汇。唯一令人沮丧的是，你完全读不懂

它们在说什么。

◎功能词：短而隐性

回过头来看一下排在前面的那20个功能词。你会发现有17个词是由3个或3个以下的字母组成的。每种语言中最常用的词往往比较短，且通常只有一个单一的、易于发音的音节。

隐性词汇不仅短小，而且难以被察觉。我们没有察觉到林肯演讲中高频使用的功能词的原因之一是我们的大脑自然地略过了它们。我们不由自主地将注意力集中在了与内容相关的词汇上。我们记忆词汇的方式也体现了功能词的隐性性质。比如说，回想一下你和某人的最后一次谈话。你能想起这个人所说的具体的词汇吗？有非常大的可能性，你记得的只有实义词。

也许对隐性最强有力的测试是主动去听一个人所使用的功能词和实义词。通过看电视、听广播，或者只是听身边人的谈话，并且有意识地去关注功能词，你会发现功能词说得非常快，持续时间平均不到1/5秒。事实上，这样的速度经常在心理学实验中用来呈现几乎无法被感知的文字或图片。想象一下你在几分钟内一直关注功能词，那么你会发现自己完全不知道谈话的内容是什么。只关注功能词本身几乎是一件不可能完成的事情。

约翰·克里的事例和他的隐性词汇

在2004年的总统竞选中，民主党人约翰·克里参加竞选，当时在任的是乔治W.布什总统。在接下来几个月的竞选中，克里对布什的支持率构成了严重威胁。然而，克里身上一个反复出现的问题是，他无意中会表现得有一点冷漠和高傲。讲话时，他的肢体语言非常僵硬且不友好。他的演讲和采访常常表现得呆板和不真实。

据《纽约时报》的一篇文章报道，为了表现得更亲民，克里的顾问建议他更多地使用“我们”，减少使用“我”。阅读到这篇文章时，可以很明显地感觉到克里遇到了麻烦。

我们后面会进行详述“我”字的使用与个人和诚实相关，政治人物使用“我们”会显得冰冷、死板和疏远。在当时，克里使用“我们”的频率已经是布什的两倍，使用“我”字的频率是布什的一半。克里的顾问，可以说是这个国家里最聪明的人之一，也没能够领会到隐性词汇是如何运作的。

这应当作为一个有意义的教训。功能词几乎是听不到的，且你对功能词的惯有想法很有可能是错的。

功能词另一个让人吃惊的方面是，人在大约12岁之后就很难掌握它们。成年后再学习另外一门语言通常是很困难的事情。然而，大部分人可以很快地学会描述物体、数字和颜色的词。他们也能够记住“为了……”“在……里面”“在……上面”“和……一起”（for、in、above、with） 等相关词汇。但是掌握正在进行的对话中最常见功能词就困难很多。事实上，通常通过一个人的写作，你就可以判断出他不是英语母语者。他们通常会在功能词上出错，而不太会在名词或者常用动词上面出错。

◎功能词和大脑

从大脑损伤的部位也能够看出功能词和实义词间的区别。有时候，中风或脑部受伤会损伤到脑部左侧一个高度特定的部位。如果是在某个区域中，这个人可能会丧失使用实义词的能力，但是仍然能够使用功能词。如果是另外一个区域则可能导致相反的结果。

这两个大脑区域——布罗卡区和韦尼克区位于大脑左侧的外表面上，被称为大脑皮层。

布罗卡区命名自19世纪的法国外科医生保罗·布罗卡，位于大脑额叶。19世纪60年代，布罗卡发表了一系列的文章说，布罗卡区受损与患者痛苦而缓慢的、不连贯的说话方式有关联。然而更令人吃惊的是，他们通常不能够有效地使用功能词。举个例子，如果一个布罗卡区受损的患者被要求描述本章开头的那幅图，他可能会说：“女孩……嗯……女性……啊……图，嗯……老。对，老妇女。”通常，布罗卡区受损的患者在社交方面很尴尬，他们常因无法与其他人正常沟通而感到沮丧。

几年后，卡尔·韦尼克发表了关于大脑颞叶的发现，这个区域被称为韦尼克区，这一发现使得布罗卡区更为有意义。韦尼克区受损会产生完全不同的症状。具体来说，韦尼克区受损会导致人无法使用名词和规则动词，与此同时，却不影响他们使用功能词。如果一个韦尼克区受损的患者被要求描述本章开头的那幅图，他可能会说：“那么，那个人就在那里，应该在她旁边。如果我往那边看，你也会看到她。我们来看。我正在想她呢。她就在那里。”

如果说布罗卡区控制着功能词，韦尼克区控制着实义词，那么这种说法过于简单化了。但是它道出了一个事实：实义词和功能词的不同在大脑中有本质的区别。



左侧视角

尤其值得注意的是，布罗卡区（与功能词相关的区域）处在大脑额叶。大脑额叶部分控制的几项技能大部分都与社交有关。比如说，已经有很多研究解释了额叶是如何与表达和隐藏情感的能力相关联的。其他额叶区与阅读他人面部表情的能力有关。许多有前景的研究表明，很多控制我们情感能力和与他人建立社会关系的能力与额叶活动相关。

也许额叶损伤和社会行为及人格变化最显著的例子是菲尼斯·盖奇的案例。盖奇是19世纪中期的一位铁路爆破专家。据说他是一个认真、尽责和严肃的人。一年夏天，一次爆炸清除岩石的准备中，盖奇需要把爆破粉夯实。夯实的过程中，他不小心在长塞药棒上摩擦出了一个火花，火花点燃了炸药，导致塞药棒向上飞了出去，正好在头骨上砸了一个洞，并摧毁了他的大部分额叶。出乎所有人的意料，盖奇并没有丧生，他在几个星期内就恢复了身体健康。除了他脑袋前面的洞。在接下来的几个月，菲尼斯·盖奇的性格发生了巨大变化。他从缄默变成聒噪，从小心谨慎变成冲动，从有礼貌变成令人憎恶，还有从清醒变成，怎么说呢，不清醒。他变成了一个完全不同的人，而且再也回不到以前的性格了。

20世纪初，巴甫洛夫在他的狗身上发现了同样的事情。巴甫洛夫是第一批诺贝尔奖的获得者之一，他被人熟知的实验是一只狗被训练为铃铛响起时就会流口水。在寻找经典性条件反射区失败的尝试中，巴甫洛夫在手

术中损坏了狗狗们大脑的不同区域。他报告说，只有损伤额叶才会影响到动物的个性。虽然额叶受损的狗还记得手术前的事情，但它却和之前不同了。

如果额叶与性格和社会行为联系紧密，那任何位于额叶的语言区（如布罗卡区）也与性格和社会行为相关联就不足为奇了。

◎功能词非常非常社会化

大脑研究得出了不可避免的结论：功能词与社交世界相关联。事实上，隐性词汇的本质就是社会化。

想象一下，在一个有风的下午，你漫步在街道上，突然有一张纸落在你前方的地面上。上面手写着：

他就在附近，但是我不知道在哪里。我很快将会回到这里。不要！

上面的话语法正确，并且在一定意义上是可以理解的。有什么重要的事情要发生吗？显然是有急事。但是，确实，我们不知道这个人在说什么。纸上的每一个词几乎都是功能词。“他”在哪儿，又是谁？“将会”和“很快”是什么时候？“我”是谁？“这里”是哪里？不要这个人做什么？现在再来想一下，这张纸上说的简直毫无意义。

在一个正常的交谈中，基于我们的聊天对象、所处位置和我们已经聊过的内容，我们会自然而然地知道所有功能词指代什么。不论是谁写的纸条，写的人和目标收件人对于人物、地点和时间有着共同的理解基础。也许写纸条的人是鲍勃，读纸条的人是茱莉亚。几分钟前，他们也许刚进行了下面的对话。

鲍勃（电话另一头是茱莉亚）：茱莉亚，我现在有点忙。我得去买一个订书机，不过我会在门口留一张纸条，以防你到的时候我不在。

茱莉亚：好的。我需要会计师签署我的费用表。你知道他在哪里吗？

鲍勃：我去看看他在不在……

茱莉亚：我有告诉过你我打算复吸了吗？每当吸烟的时候我都感觉更加警觉和快乐。但是我知道你会反感。

鲍勃：你疯了吗？我们得谈谈。现在我得走了。回头见。

一瞬间，我们就知道了“他在附近”＝会计师在大楼里，“我”＝鲍

勃，“会很快回来”＝鲍勃大概会在写完纸条的半小时内回来继续工作，“不要”＝茱莉亚不应该再次开始吸烟。

有趣的是，这个纸条的真正的含义大概只有鲍勃和茱莉亚在当天的特定情境下才能够理解。如果茱莉亚是在一周以后发现这张纸条，那也不再有意义。任何恰巧看到这个纸条的陌生人都无法破解这些功能词的真正含义。

恰当使用功能词需要社交技巧。说话者假定听者知道所说的每一个人都是谁。听话者需要了解说话者，并集中注意力才能够理解对话的内容。因此，仅仅是理解一段简单的充满功能词的对话也是需要社交知识的。

冠词、介词和所有其他隐性词汇也是如此。看一下下面这些具有细微差别的句子：

“我简直不敢相信他给了她那枚戒指。”

“我简直不敢相信他给了她一枚戒指。”

“那枚”和“一枚”之间的差别微妙而显著。如果使用了“那枚”，那么意味着说话者在指代一枚听话者也知道的特定的戒指。相反，如果使用了“一枚”，则表明“他”和“她”的关系进展了一大步。更重要的是，使用“一枚”说明说话者和听话者对这枚特定的戒指没有一个共识。

所有功能词的工作原理都相似，因为他们与说话者和听话者的个人关系相关联。即使是书的作者和读者也会共享一个社会环境。如果现在再提及前面段落所说的茱莉亚吸烟的事情，你会立刻就知道“我”“现在”“到”“早点儿”和“关于”指的是什么。如果这些词出现在三页前，没有人能够理解其中的意思。所有的功能词，如“之前”“在……上面”和“到……”，需要对说话人所处的时间和地理位置有个基本的了解。因而，使用功能词的能力是基本社交技巧的标志。另外，讨论名词和动词需要能够从文化角度理解共享类别和定义。

神奇的是，我们的大脑几乎能够立刻决定应该使用什么功能词。假设你的母语是英语，从来没有人坐下来向你解释“一个”和“这个”之间的差别。假设你正在和一个从未见过的人聊天，聊天的时候不经意间提到了屋子里的某个窗户，两分钟后，当提到“那个窗户”时，你们两个都会知道是指哪个窗户。同样，之后你们再次谈到它很干净时，也依然会记得“它”＝那个窗户。

功能词能够反应我们思考物体和事件的方式，也能粉饰我们微妙的思

考方式。我们能够很快地在介词和其他功能词（如冠词和代词）之间进行语言切换，和我们的说话速度一样快。很难想象我们停止一个对话，只是为了决定应该说“我去朋友家”“我朝朋友家走去”还是“我路过了朋友家”。听话者几乎察觉不到“去”“朝……走去”和“路过”之间的差别，但是它们有关这次出行、这位朋友或朋友的家都有一些不同的含意。

最后需要注意的一点是，我们不能轻松地掌握如何以及何时使用功能词。它们难以察觉、难以控制。我们的大脑处理功能词非常迅速和有效。一直以来，我们的大脑都在记忆最近提及的人物或物体，以便在接下来的句子中使用正确的代词和冠词。

语言之上：功能词体现文化

每一种语言都会区分“一个”和“那个”，“他”和“她”，以及“去商店”和“路过商店”。在一些语言中，这些区别由功能词来区分，在另外一些语言中，这些区别由周围的名词或动词来区分。文化不同，功能词的使用方式不同，通常功能词的使用方式也体现着文化本身。

在对功能词的研究中，我和我的学生、同事开发了多种语言版本的LIWC计算机程序，包括西班牙语、德语、阿拉伯语、意大利语、法语、俄语、荷兰语、汉语，以及其他语言。到目前为止，几乎所有我们发现的与社会和心理现象相关的英语语言连接都可以推广到其他语言。在开发多语种文本分析工具的过程中，每当我们开始一个新的语言版本时，都会遇到不同的问题。

省略代词

在一些语言中，很少使用人称代词。例如，在西班牙语中，“estoy triste”的意思是“感到失落”。通常前面不再加“我”字，因为人称代词已经暗含在动词的词形变化中。当然，说话者可以说“Yo estoy triste”，相当于“我感到失落”，只不过这一说法突出强调了“我”字。正如我们在情感章节讨论过的一样，当讲英语的人感到失落时，他们倾向于使用更多的“我”字——很显然是因为他们在关注自己。当讲西班牙语的人感到失落时，他们使用第一人称单数“yo”的频率会大大增加。

为什么有些国家的语言会省略人称代词而有些却不省略呢？一种观点认为，紧密型集体主义文化倾向于省略代词，而个人主义社会倾向于保留代词。

语言体现地位

大部分的语言构造能够体现出交谈中谁的地位更高或者谁更有声望。古英语中，我们的语言祖先区分“你”和“您”。到了18世纪后期，正式和非正式的区别逐渐消失。虽然大部分的欧洲语言依然在使用“你”的正式和非正式形式，但是很少做两者的区分。其他语言，例如日语，通过动词词形变化和其他词汇体现相对情况。的确，在不指出说话者和听话者相对情况的前提下，几乎不可能说“我和你谈到过那辆车”这样的话。

直接信息和间接信息

一些语言，如土耳其语，要求你为自己的每一个说法提供根据。如果我用英语对你说“奥斯汀昨天天气特别热”，你大概会耸耸肩，然后认为我说的是事实。但是，在土耳其，你需要使用不同的动词形式来表明你是亲身经历了炎热的天气还是只在转述这个信息。

翻译中丢失社会信息

在一系列引人注目的研究中，斯坦福大学的莱拉·博格迪特斯基展示了你使用的语言如何决定你记忆图片和事件的方式。如果用日语介绍另外三个人而不是用英语，那么日英双语者会更可能记住另外三个人的相对情况。如果分别用土耳其语和英语谈论奥斯汀的天气，那么土英双语者会以不同的方式记住这次谈论。

事实是，当一种语言被译为另一种语言时，有些内容会消失或被创造。如果我要将英文的“非常感谢你”译为西班牙语，就需要去推断应该用正式还是非正式的“你”。而当相同的话回译为英语时，正式性的信息就会消失。

有趣的是，名词和规则动词的翻译通常比较顺利。往往是功能词的翻译会出现比较大的困难。

[1] 张培基译文。——译者注

语言风格和心理：跨向下一阶段

功能词无处不在。我们一直在使用和接触功能词，但是很难听到它们和掌控它们。很多功能词涉及说话者、听话者以及他们之间的关系。但是本书真的不是关于功能词本身的。如果你和一个朋友聊天，说的是“一把椅子”而不是“这把椅子”或者“那把椅子”，那么这里面关于你的信息就非常少。但是，如果我们计算你一天或一周使用的冠词数量呢？如果我们发现一些人使用“某个”（a）和“那个”（the）的频率非常高，而另一些人几乎完全不使用冠词呢？

事实上，有些人使用冠词的频率真的非常高，而另一些人很少使用冠词。从成千上万的语言样本中发现，男性使用冠词的频率高于女性，样本包括书籍、博客，还有日常对话。而且，使用冠词频率高的人更有组织性、情感更稳定，即使将性别考虑在内情况也是一样。确实，不论男性还是女性，使用冠词频率更高的人更加谨慎，政治上更加保守，年纪也更大一些。

现在，事情开始变得有趣了。在日常交际中使用冠词不会使人成为一个适应性强、年纪大的保守政治家，比如约翰·麦凯恩（他在2008年的总统竞选中，使用冠词的频率高于他的对手）。相反，从冠词的使用中可以看出一个人的思考方式、感受方式和与其他人的沟通连接方式。代词、介词和几乎所有的功能词都是如此。

这是我文章的核心。我们可以通过倾听、计算和分析功能词来了解人们，虽然这个方式可能他们还不理解。与此同时，人们使用功能词的方式会潜移默化地影响我们对人物或信息的认知。在开始功能词和人身心状态旅程之前，你可能需要一个简要的路线图以唤醒关于不同功能词含意的记忆。在本书的最后，附有一个简短的词汇识别指南。当你在学习母语或其他语言的词汇时，可以在需要时翻看该指南，以便更好地理解功能词含意。

第3章　语言体现性别、年龄和权力

男性、年长者和较高社会阶层的人使用名词群的频率都很高；女性、年轻者和较低阶层的人使用代动词群的频率都很高。

要讨论人与人之间语言使用的差别，从性别开始讨论是最好的选择。男性和女性使用词汇的方式是否不同？你可能已经猜到了，答案是：不同。既然这个问题的答案已经确定，让我们来做一下下面的测试。

在日常对话、邮件、非正式谈话、博客，甚至是大部分的书面语中，谁会更多地使用下列形式，男性还是女性？圈出每个问题的正确答案。

1.第一人称单数（如我、我的 ）

a.女性使用的更多

b.男性使用的更多

c.男女之间没有差别

2.第一人称复数（如我们、我们的 ）

a.女性使用的更多

b.男性使用的更多

c.男女之间没有差别

3.冠词（某个、这个）

a.女性使用的更多

b.男性使用的更多

c.男女之间没有差别

4.正性情绪词（如爱、有趣、好 ）

a.女性使用的更多

b.男性使用的更多

c.男女之间没有差别

5.认知型词汇（如认为、推论、相信 ）

a.女性使用的更多

b.男性使用的更多

c.男女之间没有差别

6.社交词汇（如他们、朋友、父母 ）

a.女性使用的更多

b.男性使用的更多

c.男女之间没有差别

这个测试对每个人来说应该都很简单。每天我们都会听到男性和女性说的成千上万个字。事实上，我们所有人的一生都充斥在男性和女性的话语中。所以谁会选错这些问题呢？

然而结果是，大部分人都会选错。其中包括许多世界各地顶级心理学系和语言学系的学科带头人。如果你的高中英语老师做这些题，结果也不会好到哪去。你呢？答案是：1，a；2，c；3，b；4，c；5，a；6，a。女性使用第一人称单数、认知型词汇和社交词汇更多一些；男性使用冠词更多一些；对于第一人称单数和正性情绪词的使用，男女之间没有差别。如果你跟大多数人一样的话，那你可能选对了社交词汇这一题，然后其他题都没选对。

在进一步讨论之前，复习一下测试题的答案会有所帮助。

1.女性使用第一人称单数，或者说“我”字，比男性更多一些。代词的使用体现出人们的关注点在哪里。如果一个人在焦虑、不自在、痛苦或郁闷的时候，他们会更多地关注自己。研究表明，通常情况下，女性的自我意识比男性更强，更关注自我。男女之间在使用“我”字上的差别还是挺明显的。在日常交流、博客和演讲中，女性使用“我”字的频率会高很多。

2.男女使用第一人称复数或者“我们”的频率基本一样。“我们”是词汇表中最神秘的虚词之一。最自然的假定是，当说话者使用“我们”时，他们指的是自己和自己亲密的朋友。甚至有些著名的心理学研究让被试者在屏幕上看到闪过的“我们”字样，会让被试者感到温暖和模糊，感到与他人的

联系更紧密。

事实证明，“我们” 确实是与众不同的两个字。确实，“我们”有温暖又不分彼此的感觉，如“我的妻子和我”“我的狗和我”“我的家人”。你可以感觉到“我”和“们”像两只热情的手臂一样环抱着我们。所以，有时候“我们”确实会带来强烈的团体认同感。

然而，有时候的“我们”却是另外一种情况，是稍有一些冰冷的、有距离感的和没人情味的。我带的研究生对这种“我们”深有同感，比如在我说“你们知道的，我们真的需要分析一下这些数据”的时候。当我对我的儿子说“我们需要把垃圾扔出去”的时候，他不会感觉到温暖。因为当我在说这些的时候，我没有打算去分析数据，也没有打算让全家人都去倒垃圾。在很多情况下，人们虽然使用“我们”，但其实意思是“你”。“我们”只是一种礼貌地命令人的方式。

“我们指你”的一种冰冷的变体是皇室的“我们”。据说维多利亚女王说过“我们不开心”，其实意思是女王自己不开心。在英国首相玛格丽特·撒切尔的第一个孙子出生不久后，她向媒体宣布：“我们已经是祖母了。”王室、官员和高级管理人员偶尔会在应该使用“我”字的时候，不经意间使用了皇室的“我们”。

最后，就是受政治家们青睐的模糊的“我们”了。“我们的国家需要改变，而且我们可以做到！我们的税收过高，我们需要做些什么改变！”有时候我和我的学生围坐在一起，试图拆析政治演讲，以便弄清楚“我们”代表什么。有时，它代表“你们”，有时代表“我”，有时代表“你和我”，还有时候代表“世界上所有赞同我的人”。

之所以说“我们”是一个有意思的词，是因为一半情况下，说话者使用它是为了拉近与他人的距离，另一半情况下，说话者使用它是为了转移责任。而且确实，使用上男女有差异。女性倾向于使用温暖式的“我们”，而男性倾向于使用冰冷式的“我们”。然而通常情况下，男女使用“我们”的频率相近。

3.男性比女性更常使用冠词（某个、那个）。除了仔细阅读过本书最后一章的人，几乎没有人知道这一题的答案。冠词？谁会关注冠词的使用？我算一个。而且我个人想要关注，是因为冠词是非常重要的词汇。冠词后面接名词，尤其是具体名称和特定名词。一个人使用冠词来表示一个特定的物体或事件。男性谈论物体和事件的频率高于女性。他们会谈论坏掉的汽化器，谈论他们的妻子，谈论烤架上为晚餐准备的牛排。一会儿我们会回到这个简单的概括上。

4.男女在正性情绪词的使用上没有差别。虽然在日常谈话中，女性使用负性情绪词比男性稍微多一些，但是两性在使用正性情绪词上频率一样高。

5.女性使用认知型词汇的频率高于男性。认知型词汇反映了不同的思维方式，包括那些涉及洞察力（理解、认识、思考 ）、因果思考（因为、原因、基本原理 ）和相关领域的词。女性比男性更多地使用这些词汇这一事实给了亚里士多德一记耳光。亚里士多德认为女性没有男性理智，且不具备哲学思维。这里有了一个简单的解释。最后，社交词汇的使用成为关注的焦点。

6.女性使用社交词汇的频率远高于男性。社交词汇指任何涉及他人的词汇。相信你肯定做对这一题。女性确实会更多地想到和谈论到他人。

事实上，两性在词汇使用上的各种不同形成了一个连贯的故事。什么时候男性和女性会聚在一起，他们聊些什么呢？当女性在谈论其他人的时候，男性却谈论着毫不相干的话题，像是汽化器和其他的物体和事件。最终哪个话题（其他人和汽化器）更为复杂，需要更多的认知型词汇来解释呢？人际关系不是航天器学，而远远比航天器学复杂。我们可以把顶尖的科学家聚在一起，然后把人类送上月球。两个说话者（男性或女性）可以在一小时内解决汽化器的故障问题。但是，即使是最具有创造力和最勤奋的科学家，都没有办法理解或解释女歌星詹妮弗·洛佩茨为什么迷上了那个她迷上的男性，或者在她和现任丈夫的婚姻还会维持多久这个问题上达成一致。

◎为下一个测试做好准备：另一个区分男女的词汇类别

你刚才做的有性别差异的测试只是一个开始。男性和女性在语言的其他各个方面都体现出不同，全书将讨论这些方面的不同。假设会有另外一个性别/语言测试，那么鉴别男女在以下方面的不同是很有必要的。

男性更多使用	女性更多使用
大词	人称代词
名词	动词（包含助动词）
介词	负性情绪（尤其是焦虑）
数词	否定式（不、不是、从不）
每句单词数量更多	确定性词汇（总是、绝对）
脏话	模糊限制短语（“我认为”“我相信”）

这些附加的语言区别支持了上述已经提出的观点。男性通过对他们面对的事物进行计数、命名和整理来归类他们的世界。除了使话题个人化以外，女性谈论的话题更多变，她们关注话题的转变，讨论变化需要更多的动词。

最后，女性语言研究方面，涉及最多的是模糊限制短语或模糊限制语。模糊限制语通常以“我认为”“在我看来”“我不知道，但是……”这样形式的短语作为一个句子的开始。思考一下下面这个问题的两个答案的意思。

“外面天气什么样？”

“我觉得挺冷的。”

“挺冷的。”

“我觉得挺冷的”这个答案和“挺冷的”这个答案相比，表达出的含义不仅仅是外面的温度。“我觉得”这个词暗示着：“虽然我们对此看法不同（因为你也许会得出不同的结论），但是我的个人看法是外面可能会冷。当然，我可能会错，但是如果你跟我对温度的感受不同，我不会感到生气。”“我觉得”暗含着可以有多种看法的意思，与此同时，表达出对温度的估计只是个人看法而不是事实。如果只说“挺冷的”，就是在说外面温度低，外面冷。这是一个不可反驳的事实。讨论结束。

女性比男性更喜欢使用模糊限制语并不奇怪，这与女性的社交兴趣相符。有意思的是，女性不论是和同性还是和异性交谈时，都同样喜欢使用模糊限制语。

性别、刻板印象和计算机数据

男性和女性的语言使用方式没有什么深藏的秘密。然而即便如此，我们大部分人还是忽视了两性间的语言区别，我也是一样。在单词计数的计算机程序LIWC研发成功后不久，我就开始用它来分析男性和女性写的文章、博客和其他文本样本，分析数量从几十到上百再到上千。第一批的分析结果之一就是揭示了刚才文中讲到的效果模式。第一次发现女性比男性更多使用“我”字时，我直接忽略了结果。我认为肯定是偶然，再次做了研究，同样的结果。又一次的偶然，我想。大概在使用大量样本做了十几次的分析后，我才恍然大悟。

我们对男性和女性的刻板印象是根深蒂固的。即使是在科学界，对两性间语言区别的研究也是高度政治化的。一部分科学家强烈认为，男性和女性在本质上是相同的；另一部分科学家认为，男性和女性是完全不同的。其他人只是不想考虑这个问题。但是对两性的刻板印象依然存留。有些人认为女性更情绪化，男性更具逻辑性，或者认为女性爱谈论别人而男性更倾向于谈论自己。有些人相信女性每天说的话是男性的两倍。虽然很多优秀的科学研究已经揭开了这些结论的真相，但是这样的印象还是难以摆脱。

女性运动开始于20世纪60年代末70年代初，这使人们意识到了男女不平等的事实，并在文化领域有所体现。加州大学伯克利分校的语言学家罗宾·洛克夫在1975年出版了令人感到震撼的书——《语言和女性地位》（Language and Women’Place），书中指出了男性和女性在说话方面的不同之处。洛克夫指出，男性使用的语言粗鲁，体现权力，而女性讲话通常更平和、被动且过分礼貌。在接下来的几十年里，有几项研究支持了拉科夫的发现。1990年，乔治城大学的语言学家黛伯拉·坦纳用她的《男女亲密对话》（You Just Don’t Understand）引起了世界的关注。这本书连续四年位于《纽约时报》的畅销书榜单上，书中指出男性和女性在交谈时往往各说各话，完全意识不到不同性别就犹如不同文化一样。比如说，女性非常关注别人的想法和感受，而男性欠缺一些。女性认为男性说话生硬且冷漠，男性认为女性说话隐晦且不直接。

洛克夫和坦纳等社会语言学家关注的维度很广泛，如性别、种族、社会阶层和权力。他们的方法是定性的，包括逐案地记录和分析对话。这是一个缓慢而费力的工作。这样一年下来，一个优秀的社会语言学家可能也只能分析几段对话。虽然定性分析在对一小部分对话进行深入理解方面比较有效，但是这个方法不是用来对整个社会或文化进行精准描述的。这时就需要基于计算机的文本分析方法的帮助。比如，通过分析上万人的博

客，基于计算机的分析方法可以根据年龄、地位、母语、地区和其他影响因素很快确定两性区别的本质。换句话说，进度相对较慢但是分析仔细的定性方法可以更深入地了解一小群人，基于计算机的定量方法提供给我们更为广阔的社会和文化视角。这两种方法虽然互补，但是使用这两种方法的研究阵营却互不理解。

两性间的语言差异有多大

虽然男性和女性使用语言的方式不同，差别却是微妙的。通过对超过1.4万条语言样本的研究，我们发现人称代词在女性话语中占14.2%，相比之下，在男性话语中只占12.7%。从数据角度来看，这是一个很大的差别。这种巨大的数据差别能使科学家产生喜悦的泪水（至少是我）。

但是数据可以产生这种催泪效果，在生活中却不总是能产生有意义的效果。比如说，一个人说话很慢，每分钟讲100个词。在一分钟之内，一个女性使用14.2个人称代词，一个男性使用12.7个。也就是说，一个女性每分钟比男性多使用1.5个人称代词。把这些数字按天加起来。比如说，每个人平均每天讲1.6万个单词。一年下来，一个女性会比一个男性多说8.5万个人称代词。这使我感到很惊奇。但是又有什么关系呢？虽然女性比男性使用人称代词的频率高出近12%，但是我们的大脑无法捕捉到这一差别。一般人不会意识到这12%的差距，甚至摆在书面上也察觉不到。如果打算做隐性词汇检测的话，你需要人工一个个去数，或者让电脑来做这项繁杂的工作。

我们也许不擅长捕捉两性间语言上的差别，那么计算机在区分两性的文字方面做得如何呢？假设我们存储了10万篇微博文本，这些文本来自，嗯，我不确定，就当是来自19320个能够确认作者性别的微博里。想象一下，这之后，开发我们的计算机程序，使其能够根据虚词的使用将这些博文按照作者性别进行分类。正如你猜想的那样，我们已经这么做了。最终结果是，计算机的正确率为72%（正确答案的概率是50%）。如果我们把写作的内容（除了语言风格外）也包含进去的话，正确率略有提高，为76%。值得注意的是，计算机的正确率远高于人的猜测，人的正确率为55%～65%。

所有这些数据告诉我们，通常情况下，男性和女性使用语言的方式不同。而这些不同是微妙的，人类不易察觉。当然，这之间的不同远比我所提到的要复杂得多。所有人根据所处情境的不同，都会改变他们的说话方式。比如说在正式的场合，人们倾向于少用代词和社交词汇，多用冠词，这意味着大多数人在正式场合讲话像典型的男性。当和家人在轻松的环境中闲聊时，我们讲话更像女性。总之，语言中存在的性别差异能够在一定程度上反映出性别。但是说话的语境对语言的选择影响可能更大一些。

对人物性别的把握：剧作家和编剧

作为读者，你也许对文字情有独钟。你可以很容易地看出哪篇文章写得好，哪篇文章写得不好。作为一名老师，我一直在想，一个人如何成为一个优秀的作家。更进一步，如何成为一个伟大的作家。小说家史蒂芬·金写过一本令人愉快的书——《写作这回事》（On Writing），他在书中讲到，一个有能力的作家通过练习和努力有可能成为一个优秀的作家。但是，一个优秀的作家无论通过什么样的训练都无法提升为伟大的作家。史蒂芬·金认为伟大的作家是与众不同的类型。

写作文学领域常见的主题是，一个伟大的作家在语言方面有惊人的天赋。捕捉对话的能力类似于音乐方面有绝对音感。世界级对白作家似乎天生就知道如何通过语言去塑造形形色色的人物。想一下那些话剧和电影领域的著名对白作家：阿瑟·米勒的《推销员之死》（Death of a Saleman）中鲜明的语言，约瑟夫·曼凯维奇的《彗星美人》（All About Eve），以及莎士比亚的几乎所有作品，汤姆·斯托帕德、伍迪·艾伦、托尼·库什纳和诺拉·依弗朗。

几年前，在我的一次关于性别和语言的演讲中，有一个人问道，是不是剧作家和编剧天生就知道男性和女性分别是如何讲话的。恰巧我的妻子是一名很优秀的专业作家，我回想了一下她的作品，然后非常确定地答道：是的。虽然据我所知还没有相关研究，但是在赌金额度合理的情况下我敢打赌，那些天生就特别擅长对白的人一定会使用适合他们人物性别的语言。专家也是这么说的。

那一晚我在酒店房间里查了一下我的电脑，看看有什么剧本或电影脚本可以用来分析。我分析了莎士比亚的《罗密欧与朱丽叶》（Romeo and Juliet）和诺拉·依弗朗的《西雅图夜未眠》（Sleepless in Seattle）。两部作品都有一个男主角和一个女主角，一部作品是男作者，另一部是女作者。在几百次的键盘敲击之后，我有了答案。罗密欧先生和朱丽叶小姐说起话来都像男性，而依弗朗的男女主角说起话来都更像女性。用阿道司·赫胥黎的话说，就是美好的假设被丑陋的事实扼杀。但是，也许这两位作者或者这两个剧本只是巧合。

几天以后，我遇到了我的研究生学生莫莉·艾尔兰。虽然莫莉非常了解心理学，但是本科阶段对文学和哲学方面的培养使她对戏剧和电影有着格外广阔的视野。她非常热切地承担了一个大型项目，分析来自70多个不同的剧作家和编剧的110个剧本。这样的项目可以解答下面几个问题。作者在捕捉男女语言方面的能力是不同的吗？是否捕捉到的两性语言的不同取

决于特定的戏剧或电影？也就是说，你可以想象一下，在一部电影里，作者笔下一个敏感的男主角说话像个女性，而在另一部电影里像个铁血猛男。

幸运的是，多年来我们实施了几项研究，研究中我们让上百人在日常生活中随身携带录音机。我们还收集了来自面试、实验研究以及其他来源的语言样本。这些记录为我们提供了男女间成千上万次的对话。利用这些数据，我们可以将现实生活中男性和女性的说话方式和文学作品中男女主角的说话方式进行对比。

结果是令人着迷的。为了使结果变得稍微简单些，我和莫莉把数据统计结果转换成了1～9的男性化-女性化语言等级。1分表示主角说话像个十足的男性。你知道十足的男性是怎样说话的——很多冠词和介词，非常少的代词、社交词汇和认知词汇。9分表示人物说话像个十足的女性——说话模式基本上和十足的男性相反。5分表示人物的语言不体现其性别。

男女主角的男性化－女性化语言等级

1 2 3 4	5	6 7 8 9
极具男性化特征 语言	**不体现性别特征** 既不男性化也不女性化	**极具女性化特征** 语言

女性说话像男性，男性说话像女性的著作

作者	作品	女主角	男主角
琼·图克斯伯里	《纳什维尔》	9	4
斯派克·李	《为所应为》	8	5
大卫·林奇	《蓝丝绒》《穆赫兰道》	6	4
山姆·谢泼德	《被埋葬的孩子》《疯狗布鲁斯》	6	4
桑顿·怀尔德	《我们的小镇》	6	1

男女说话都像女性

作者	作品	女主角	男主角
诺拉·依弗朗	《电子情书》《西雅图夜未眠》	9	6
格特鲁德·斯泰因	《布鲁斯与威利》《不是姐妹的三姐妹》	9	8
索菲亚·科波拉	《迷失东京》	7	8
伍迪·艾伦	《汉娜姐妹》	6	7
卡莉·克里	《末路狂花》	6	8

男女说话都像男性

作者	作品	女主角	男主角
昆汀·塔伦蒂诺	《低俗小说》	2	4
莎士比亚	《罗密欧与朱丽叶》《泰特斯·安特洛尼克斯》	3	3
洛纳·威廉姆斯	《美丽比一比》	3	3
卡梅伦·克罗	《几近成名》《甜心先生》	4	3
科特妮·亨特	《冰冻之河》	5	3

注：数字1～4表示语言更倾向男性化，6～9表示语言更倾向于女性化。

以上表格给出不同的作者描绘他们笔下人物的不同方式。桑顿·怀尔德在《我们的小镇》（Our Town）中记述了几个人物的生活轨迹，包括艾米丽·韦伯。我们知道艾米丽是一个热情的高中生，她后来爱上了乔。在死于分娩后，艾米丽最后一次出现在墓地，她刚从童年的场景中及时赶回。最后一幕，当她回到自己的墓地后，艾米丽和其他亡灵谈论起她看到自己早年家庭生活时的震惊感。

艾米丽： 我从未意识到。所有的事情就这么发生着，而我们从来没有注意过。带我回去——在山丘上——回我的墓地。但是首先，等一下！让我再看一眼……有人曾注意过自己生活的每一分每一秒吗？

舞台监督： 没有。圣人和诗人，也许他们偶尔会留意生活……

艾米丽： 噢，斯廷森先生，我应该听他们的话的。

西蒙·斯廷森： 是的，现在你知道了。现在你知道了！这就是活着的样子。在无知的行云中走来走去；来来去去践踏感受……你的感受……消磨和浪费时间，就好像你有100万年的时间。永远以自我为中心，或以另

一个人为中心。现在你知道了，这就是你想要回去的幸福生活。无知和盲目……

艾米丽： 他们不明白，是吗？

茱莉亚·吉布斯： 是的，亲爱的。他们不明白……

舞台监督： 几乎所有格罗佛角的人都睡着了。几乎所有的灯都熄了——肖蒂·霍金斯在停车场，刚看着一辆奥尔巴尼火车从这里经过。马房中有人在整理和谈话。嗯，都收拾干净了。星星在空中做古老的十字交叉旅行。学者还有问题未解决，但是他们似乎认为上面没有生物。只有白垩岩……或者火。只有这一个在过滤，过滤掉所有时间来取得成功。过滤的结果非常失败，每过16个小时，所有人就要躺下休息。嗯……现在是格罗佛角时间晚上11点。你也好好休息。晚安。

如果你没有读过《我们的小镇》，我可以告诉你这是一本著作。即使在这短暂的一幕中，女性——艾米丽和她的婆婆茱莉亚·吉布斯关注着她们自己和其他人的感受。男性——西蒙·斯廷森和舞台监督几乎不使用代词，以男性的方式客观地描述这个世界。怀尔德笔下的男性的确说起话来就像典型的男性，女性就像典型的女性。

把怀尔德的作品和卡莉·克里的《末路狂花》（Thelma and Louise）做一下对比。

经历了一系列难以忘怀的不幸后，塞尔玛与露易斯（剧中的两个主角）开始逃避枪杀未遂强奸犯的法律制裁。首要调查员哈尔通过电话请求露易斯放弃逃亡。

露易斯： 如果我告诉你这一切都是意外，你会相信吗？

哈尔： 我相信你。我也想让大家都相信。可问题是，这一切看起来不像是意外，而你又不当面告诉我这一切……我需要你的帮助……你能过来一趟吗？

露易斯： 我不过去。

哈尔： 那我只能抱歉了。我们将以谋杀罪起诉你。现在，你想要活命吗？

露易斯： 你知道吗，一些词汇和字眼刚刚从我脑中飘过，像是禁闭、体腔检查、无期徒刑、电击死亡之类的词。所以，活命？我不知道。让我们想想。

哈尔： 露易斯，我会尽我所能。我知道你为什么逃亡。我知道你在得克萨斯发生了什么事。

克里笔下的男性更有趣的原因是他们说起话来比女性还女性。哈尔使用代词频率很高，使用冠词频率较低，表明他对其他人而不是对具体对象有浓厚兴趣。这部电影的另外一个重要特点是剧中所有男性（包括布拉德·皮特饰演的角色）说话都是女性风格。同时，没有谁的行为会被认为是有些女性气的。

与此相反的是昆汀·塔伦蒂诺的《低俗小说》（Pulp Fiction）。剧中最传统的女性角色应该是菲比安，她和布奇（布鲁斯·威利斯饰演）是一对。尽管菲比安的长相和声音都很像女性，但是她的言语却很男性化。在这一幕中，布奇刚结束一场拳击比赛回到家，比赛中他不小心打死了对手（塔伦蒂诺笔下人物的生活都不容易）。菲比安还不知道发生了什么，跟布奇聊起了她要有大肚子的事情。

菲比安： 大肚子。大肚子还挺性感的。

布奇： 那你应该感到高兴，因为你有大肚子了。

菲比安： 闭嘴！我肚子才不大！我只是有小肚腩，就像麦当娜唱“Lucky Star”时候也有，跟大肚子可不是一回事。

布奇： 我没觉得大肚子和小肚腩有什么区别。

菲比安： 区别很大啊。

布奇： 你希望我有大肚子吗？

菲比安： 不要。男性大肚子看起来有点畸形，或者像一只大猩猩。但是女性有大肚子就会非常性感。你身体的其他部分都是正常的。正常的脸，正常的腿，正常的臀部，但是有一个大的完美的圆肚子。如果我有一个的话，我会穿小两号的T恤突显它。

布奇： 你觉得男性会觉得那样很有魅力吗？

菲比安： 我才不关心你们男性觉得什么有魅力。很不幸，女性觉得摸着舒服或看着好看的东西，男性基本上都不喜欢。

如果你经常看电影，那么昆汀·塔伦蒂诺笔下的人物使用男性语言风格可能不会让你感到震惊。但是莎士比亚这样的游吟诗人呢？下面可能是《罗密欧与朱丽叶》中最著名的一幕，两个年轻的爱人宣称如下。

罗密欧： 那边窗子里亮起来的是什么光？那就是东方，朱丽叶就是太阳！起来吧，美丽的太阳！赶走那妒忌的月亮，她因为她的女弟子比她美得多，已经气得面色惨白了。既然她这样妒忌着你，你不要忠于她吧；脱下她给你的这一身惨绿色的贞女的道服，它是只配给愚人穿的。那是我的意中人；啊！那是我的爱；唉，但愿她知道我在爱着她！她欲言又止，可是她的眼睛已经道出了她的心事。待我去回答她吧；不，我不要太鲁莽，她不是对我说话。天上两颗最灿烂的星，因为有事他去，请求她的眼睛替代它们在空中闪耀。要是她的眼睛变成了天上的星，天上的星变成了她的眼睛，那便怎样呢？她脸上的光辉会掩盖了星星的明亮，正像灯光在朝阳下黯然失色一样；在天上的她的眼睛，会在太空中大放光明，使鸟儿误认为黑夜已经过去而唱出它们的歌声。瞧！她用纤手托住了脸，那姿态是多么美妙！啊，但愿我是那一只手上的手套，好让我亲一亲她脸上的香泽！

朱丽叶： 罗密欧啊，罗密欧！为什么你偏偏是罗密欧呢？否认你的父亲，抛弃你的姓名吧；也许你不愿意这样做，那么只要你宣誓做我的爱人，我也不愿再姓凯普莱特了……只有你的名字才是我的仇敌；你即使不姓蒙太古，仍然是这样的一个你。姓不姓蒙太古又有什么关系呢？它又不是手，又不是脚，又不是手臂，又不是脸，又不是身体上任何其他的部分。啊！换一个姓名吧！姓名本来是没有意义的；我们叫作玫瑰的这一种花，要是换了个名字，它的香味还是同样的芬芳；罗密欧要是换了别的名字，他的可爱的完美也绝不会有丝毫改变。罗密欧，抛弃了你的名字吧；我愿意把我整个的心灵，赔偿你这一个身外的空名。 [1]

我们能从中感觉到罗密欧和朱丽叶对彼此的渴望，能够感受到他们的单纯、明亮和率真。如果只是粗略的阅读，你不会发现两个人（尤其是朱丽叶）表达方式都偏向男性化。罗密欧和朱丽叶都较少使用“我”，使用人称代词的频率低于平均水平，冠词高于平均水平，尤其是在这样一个私密、亲密的背景下。

莎士比亚和塔伦蒂诺都是男性，笔下的人物也都像男性。无论男女角色使用功能词的方式都像男性。这两位作者使用功能词的方式可能相同，但是写作内容和写作范围完全不同。莎士比亚很有趣，因为他能够巧妙地传达女性的真实生活和女性关心的问题。但是他使用功能词的方式和塔伦蒂诺很像，这说明他没有能够走进女性的心里。

这些从读者角度来看有关系吗？有，只不过以一种微妙的方式。回忆一下黛博拉·坦纳将性别间差异与文化间差异所做的类比。我们大部分人都读过，比如说陀思妥耶夫斯基的《罪与罚》，并对拉斯柯尔尼科夫谋杀当铺老板的残忍想法表示赞同。我们即使在不说俄语或不了解19世纪圣彼得堡文化的情况下都会这样做。当然，圣彼得堡当地的人如果在1866年阅读这本书，毫无疑问，他们会比21世纪的英语母语者对这本书的理解更为深

刻。同样的道理，一个女性作者笔下的男性角色的说话方式也会错误地刻画男性的真正说话方式，但是如果内容非常吸引人，我们可能就会忽略方式。

[1] 此处采用张培基译本。——译者注

语言中的性别差异：男性荷尔蒙的作用

男性和女性使用语言的方式不同，倾向于谈论不同的话题，最终可能代表人类的两个子文化。显然，男孩和女孩在社交中非常不同，很可能是因为语言的不同。有没有可能我们的兴趣会受到荷尔蒙的影响？如果我们的兴趣与荷尔蒙相关，那也许荷尔蒙会影响我们日常语言的使用。

男性和女性除了生理结构上的不同外，荷尔蒙的分泌也有显著不同。与男性相比，女性具有非常高的雌性激素，用以调节月经周期，促进女性生殖器、胸部和阴毛的生长，以及辅助怀孕。对于男性来说，睾丸素与男性生殖器和第二性征（如面部毛发）的发展有关。女性也会分泌睾丸素，但是分泌量会远低于男性，正如男性体内所含雌性激素也很低。

由于各种原因，男性和女性偶尔都会接受睾丸素治疗，治疗中会定期给他们注射睾丸素。当体内睾丸素含量高低不同时，他们的语言会有何不同呢？通过一系列奇怪的事件，我能够回答这个问题了。

2000年附近，詹姆斯·达布斯联系到我，他是睾丸素与心理学相关性方面的世界知名专家。他是一个礼貌、庄重、杰出的高个子美国南方人，眼中总是闪着光，詹姆斯曾测量过杀人犯、律师、演员、牧师、学者和其他人的睾丸素水平。他的大部分工作都在研究攻击性行为与睾丸素水平之间的联系。（在一个严肃的关于荷尔蒙与行为的会议上，詹姆斯环顾拥挤的演讲大厅，然后开始了他的演讲："这整个房间里的睾丸素都不足以抢劫一个贩酒的小店。"）

詹姆斯曾收到一封来自生物学上讲是女性的来信，此人正在经历女性到男性的变性治疗，或者更准确地说，是性别重置。这个人，暂且叫他GH，当时28岁，性别重置已经进行了3年。除了双乳房切除术外，GH每两到四周接受一次睾丸素注射。GH知道了詹姆斯的研究后，想知道他是否愿意研究自己保留了多年的日记。詹姆斯知道我研究语言的事情，他认为我可以检验一下睾丸素对功能词使用的影响。

GH是一个有文采和高产的作者。他为我的研究提供了自己两年来写的日记，以及他的睾丸素注射记录。当我们正在转写GH日记的时候，我被邀请去波士顿做一系列的演讲。有一晚，我在酒吧和一个绅士聊天，他正好是一位人类学家。在我们聊天的过程中，我提到了GH的研究项目。我的人类学新朋友沉默了一会儿，然后跟我说，在主治医师建议下，他也在接受睾丸素注射。他当时60岁，已经接受睾丸素注射4年了，目的是想恢复上肢力量。"真的吗？"我故作天真地问他。"那你有注射记录

吗？”有。“那你会不会恰巧也记日记呢？”没有，但是他很乐意让我分析他之前发送的所有邮件内容。

分析过程简单明了。GH和人类学家的文字都通过LIWC程序来分析。词汇类别根据每个人上次注射睾丸素后的天数进行对比。这一想法的依据是随着上一次睾丸素注射的时间推移，每个人体内的睾丸素含量会下降。

两个人结果的相似程度惊人的一致，整体结果非常有前景。不论睾丸素含量上升或下降，他们使用冠词、介词、名词、动词和消极情绪词汇的方式都没有发生预想中的变化。但是，存在一个迷人且可靠的差异——社会代词（例如我、我们、他、她、他们）。随着睾丸素含量的下降，他们使用社会代词的数量增多。考虑下这意味着什么：GH和人类学家注射了睾丸素以后，他们关注任务、目标、事件和特殊场合的物体——但不是人。几天之后，他们每天醒来开始注意身边的人，并且发现这些人非常有趣，值得聊一聊天。几周之后，另外一次注射，瞧！他们被治愈了。他们不再担心其他人了。

这项研究让我们对睾丸素对人心理上的影响有一个初步的了解。在得到分析结果之前，我询问了两位被试者，问他们感觉睾丸素是如何影响语言的。GH非常确信地说，注射睾丸素会给他更多能量，心情更积极向上，会激发性欲、攻击性、运动欲和传统的男性东西。而另一方面，人类学家坚定地认为睾丸素对自己的心理没有任何影响。两个人都错了。

这样的研究能够告诉我们两性之间的区别和功能词的使用区别是什么吗？恐怕只有一点点。睾丸素注射没有改变GH的性别和人类学家的阳刚之气。没有证据显示，这样的荷尔蒙冲击对两个人理解、思考世界或对世界加以分类的方式有任何影响。两人在冠词、名词、动词、认知词汇或介词的使用上没有一点改变，这一点很重要。最后，这项研究建立在两个不寻常的案例上，也就是说，我们在下结论时要谨慎些。

但是，从在更广的层面上讲，睾丸素研究和各种性别-语言项目告诉我们一些关于男性和女性看待世界和理解世界的方式。男性对思考和讨论他人的兴趣不如女性。这一结果可能受睾丸素的影响，但坦率地说，睾丸素结论太过初级，所以还无法做出伟大声明。

男女在社会兴趣方面有差异这一事实也不算是新闻。然而，男性一贯使用更多的冠词、名词和介词这一事实却算是新闻了。好吧，也许不能成为CNN的头条新闻，但之所以算是新闻，因为这些男女间的语言差异表明了男性更倾向于谈论和思考具体的事物和事件，并且表现方式非常具体。他们很自然地将事物进行了分类。

此外，介词的使用表明这一分类过程是以空间和分层的方式进行的。想想我们使用介词的情况。

玛利亚的汽车后备厢的钥匙放在由你母亲画的那幅船的画下面的台灯旁边。

“在”“的”“旁边”“下面”和“由”这些词详细说明了属于谁的什么钥匙放在了什么地方。从层次结构的意义上讲，这不是任意的一把钥匙。相反，它是后备厢类别的一部分，后备厢又是汽车分类的一部分，汽车属于玛利亚。同样，钥匙的位置也从水平（旁边）和垂直（下面）两个平面进行了具体的描述。据统计，这样的一个句子更可能出自一个男性口而非女性。不是男性偏爱介词和名词，而是男性天生比女性更喜欢对事物进行分类并指定物体空间的位置。

语言的智慧：一生中使用的语言

大多数大型调查中，研究人员会问三个问题：性别、年龄和社会地位。我们后续会继续讨论社会地位，这通常会通过种族、教育或收入进行间接询问。研究你的政治倾向或购买行为时，调查制作者为什么想要知道你的性别、年龄和社会地位呢？因为这三个变量能够预测出很多内容。如果我知道了你的性别、年龄和社会地位，就可以出奇准确地猜出你喜欢的电影和音乐类型，你的宗教观和政治观，你的身体和心理健康状况，甚至是你的预期寿命。

这三个人口统计数据也和语言使用相关联。在很多方面，语言和年龄之间的联系比语言和性别之间的联系更有趣。如果你天生是一个女性，你这一生都是个女性的概率是非常非常高的。然而，如果你是一个婴儿，那你不会持续成为一个婴儿的概率是极大的。

我们的语言（更具体地说，是我们使用的功能词）在一生中的改变，从某种程度上讲是非常惊人的。如果考虑了我们的目标和生活状况是如何发展的，变化就没那么惊人了，我们的身体也是一样。随着年龄增大，我们会随着朋友、家庭、性、金钱、健康、死亡和很多其他生活上的东西而改变。

我们的性格也会变化。通过对上千人的研究发现，到了大概12岁，我们的自我感觉通常才会是积极向上的。从13到20岁，我们的自我感觉下降到人生的至低点。之后，大概到70岁，大多数人的自我感觉会逐渐提升。事实上，在65岁左右，多数人对自己的感觉如同9岁时一样好——比以往都好。然后，自我感觉在生命中的最后几年会下降。我们的研究基于5400对双胞胎，研究发现，随着年龄的增长，他们变得没那么外向，情绪上更加稳定，更加受感情驱使。

性格研究与我们对老年人的刻板印象不一致，我们通常认为老人孤独、自私、古板和尖刻。一些更有前景的研究得出的结论与这些刻板形象相反，这些研究由劳拉·卡斯滕森和她的斯坦福大学同事进行。她发现人们年龄增大，情绪成为他们生命中重要的一部分。人们更加关注自己的情绪状态，他们学习如何更有效地管理自己的情绪，以获得更多的快乐并减少一些负面情绪。尽管人们在70岁以后朋友会减少，但是他们的关系网会更稳定。

法雷利兄弟的经典电影《阿呆与阿瓜》（Dumb and Dumber）中的主角劳埃德说明了一切。在老妇人看管他的物品几分钟后，劳埃德说

道："谢谢。嘿，我想他们是对的。虽然老年人又慢又危险，但还是有用的。我很快就会回来。可不要这时候死掉！"那个老妇人高兴地拿走了劳埃德的所有东西。

随着年龄增大，人们的语言如何改变呢？我们用不同的方法进行了测试。第一种方式是回到大型博客项目，我们分析了1.9万个博客使用者的帖子。我们在2004年"收获"的博客网站主要由20岁左右的人群构成。我们将样本分为三个组：青少年（年龄为13～17岁），年轻人（年龄为23～27岁）和成人（年龄为33～47岁）。忽略年龄限制，这三组人使用情感词汇和功能词的方式非常不同。青少年使用人称代词（我，你）、短词和助动词的频率非常高。年纪越大的人越倾向于使用大词、介词和冠词。

我们还做了一个系统的分析，被分析的人参与过一次我们的表达性写作实验。回忆一下第一章的内容，我长期从事研究，研究需要人们记述他们深刻的个人经历，通常是创伤经历。这些年来，世界各地实验室的研究人员会给我寄他们研究中用到的写作样本。为了这个关于年龄的研究，我和洛里·斯通·韩德尔曼合作，她是一位优秀的研究生，后来成为纽约有名的图书编辑。洛里和我分析了来自17个大学、3200个人的数据，他们都参加过45次写作研究中的一次。虽然实验中人的平均年龄在24岁左右，但是年龄范围由8到80岁。

你也许能想象得到，人们写作的主题多种多样，从性虐、毒瘾到宠物死去，以及没有进入高校足球队或者啦啦队。大多数的文章是令人心碎的。大多数人把心灵注入了他们的故事中。

计算机分析文本的结果有时令人出乎意料。在情感话题的写作中，年轻人和年长者使用词汇的方式非常不同。

不过更令人印象深刻的是年轻作者和年长作者的情感基调。与调查结果一致，年长作者使用更多的积极词汇，年轻作者更常表达消极情绪。这些区别在40岁时变得明显，在年纪大的群体中暴涨。

年轻作者	年长作者
人称代词（尤其是"我"）	冠词、名词、介词
时间指示代词	大词
过去时态动词	将来时态动词
	认知词汇（洞察词汇）

使结果更吸引人的是，所有人都被要求记录他们过往生活中最苦恼的经历。年轻一些的人会挖掘出很多暗黑词汇来表达痛苦。随着作者年龄的增长，他们使用的消极情绪词汇减少，而积极情绪词汇猛然上涨。正如下方所示，样本中最年轻的作者是一位8岁的儿童，他进入他的情感故事主题（兰迪）的方式与年纪大的作者完全不同，年纪大的这位作者正在经历癌症。

8岁的小学生： 我的敌人是兰迪，他快把我逼疯了。在外面，他叫我的名字，他不理我，不停地烦我，所以当他开始这样时我就叫他的名字。他还是不会停止，所以我就反过来这样对他，他让我发疯，我也让他发疯，所以我也不理他，但是他还是快把我逼疯了……我妈妈说他是不好的影响。

80岁的退休人员： 我今年80岁，但是我努力让自己保持忙碌和生活，仿佛这一切都不会停止。确实我不如以前那样行动灵活了，但是我不觉得这是癌症的影响，我还是尽可能做到最好。当消极情绪试图悄悄走进我的思绪时，我会立刻试着用积极的情绪替代它，想想自己还能活着是多么幸运。38次的治疗，每一次我都会早上独自开车，然后有机会做冥想，也是这时候，我从大自然的美丽和四季变换中找到了积极情绪。

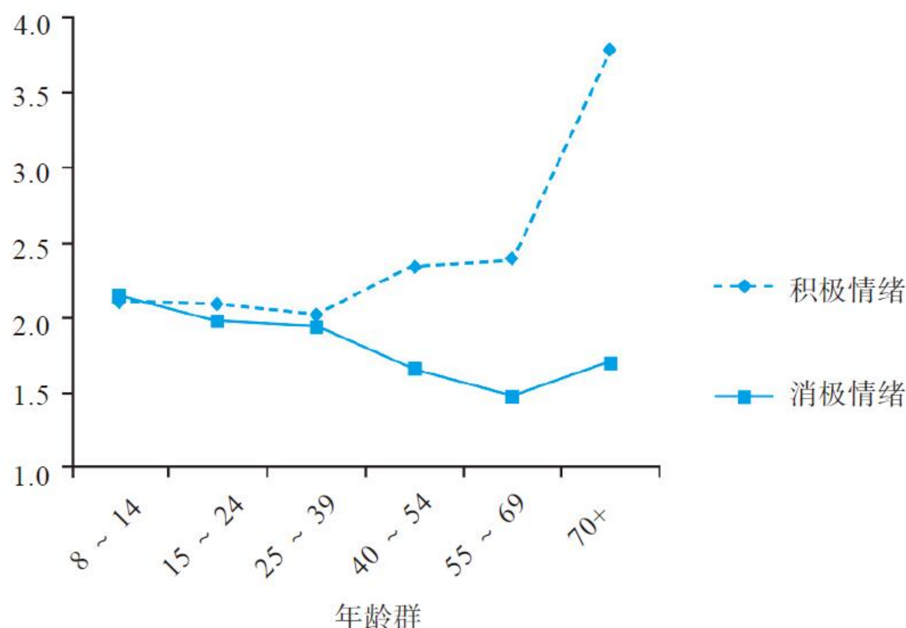


图3-1

你更愿意是80岁还是8岁？读完这些文章后，答案可能就没有你想象得那么简单了。你会对劳拉·卡斯滕森的论点更加了解：随着年纪的增长，我们能更好地管理自己的情绪。与此同时，年龄增长能让我们以更超然的方式看待这个世界。

很多寿命科学家有一个担忧是，语言的使用差异在，比如说，70岁的群体中和40岁的群体中，可能不是因年龄而产生。而是，所有70岁的人经历过同样的事情是那些40岁的人没有经历过的。例如，有些研究是在20世纪90年代做的，当时所有70岁的人在成长过程中都没有电视，且都经历了第二次世界大战。那些40岁的人通常出生时就有电话，且都是生育高峰一代的一部分。也许70岁的人乐观只是因为他们成长过程中没有电视。也许不是，谁知道呢。

我们绕过这个问题的方法之一是研究过去4个世纪的10位小说家、10位诗人和10位编剧的作品，他们一生的作品很广泛。作家这一组，我们只研究随着他们年纪增长，语言使用有何变化。总的来说，10个作家里面有8个表现出和其他项目中同样的年龄相关语言模式（另两个例外是路易莎作家梅·阿尔科特和查尔斯·狄更斯）。

一个很好的例子出自英国小说家简·奥斯汀。奥斯汀出生于1775年，去世时42岁。她开始写作时才12岁，之后一直持续写作到去世。她的原稿包括短篇小说、书信还有诗，这些都收录在后来出版的《少年读物》（Juvenilia）中；她的最后一本书《桑迪顿》（Sanditon）没能在去世前全部完成。对比一下作家简·奥斯汀第一部和最后一部作品的第一段。

《杰克和爱丽丝》——出自《少年读物》的一部小说

约翰逊先生曾经大约53岁，12个月后他54了，他非常高兴，所以决定下一个生日举办化装舞会来庆祝，舞会邀请他的孩子们和朋友们。于是，在那一天，为了这个目的，他把第55年的年票发送给他的邻居。他认识的人，在那个世界确实不是很多，其实只有威廉姆斯女士、琼斯夫妇、查尔斯作家亚当和辛普森三姐妹，他们是所有居住在帕米帝朵的人，也是参加舞会的所有人。

《桑迪顿》

一位从坦布里奇旅行回来的绅士和一位女士去往位于黑斯廷斯和边界之间的苏塞克斯海岸，因为公务原因他们放弃走高速路，而是选择了非常难走的小路，小路曲曲折折，路又长又陡，路上一半是石头一半是沙子……事故正好发生在唯一的绅士家附近的路上……一幢房子，他们的司机第一次被要求开往那个方向，司机认为那一定是他们的目标，然后带着

最不情愿地表情被迫错过了。

即使才12岁（也许那是她已有15岁），奥斯汀已经有些老成。即便如此，写作还是暴露了作者的年龄。在《桑迪顿》中，奥斯汀使用更多的介词、名词和认知词汇（例如放弃、曲曲折折、被迫），年轻时的奥斯汀使用更多的人称代词和时间指示代词（如月、天）。虽然奥斯汀不论在哪个年龄阶段使用“大词”数量都是无规律的，但是很明显，年轻的奥斯汀远没有年纪大的奥斯汀思维复杂。

奥斯汀的年龄-相关语言变化映射了那些诗人——华兹华斯、叶芝、罗伯特·格雷夫斯和埃德娜·文森特·默蕾，以及同类小说家乔治·艾略特和剧作家乔安娜·贝利。莎士比亚的情况稍微有点复杂，但即便如此他还是显示了相同的模式。

关于语言在年龄和性别上的差异还有最后一条有趣的发现：你可能已经注意到，年纪大的人通常使用功能词的方式像男性，年轻人的使用方式像女性。这不是什么统计学上的巧合。这些模式在跨文化、跨语种、跨世纪的情况下依然存在。有趣的是，不是说女性开始说话像男性，而男性保持不变。而是男性和女性通常同时发生变化。举例来说，8～14岁的人中，女孩使用词汇中代词占19%，男孩使用词汇中代词占17%。70岁的人中，女性使用词汇中代词比例降为15%，男性使用词汇中代词比例降为12%。

在讨论完社会地位差异后，我们再来讨论这些模式产生的可能原因，因为你会发现，社会地位的影响和性别、年龄的影响会有重叠。

社会阶层和语言

从一开始上学，我就被教育说美国没有社会阶层之分。当然，英国和印度有，但是美国没有。在美国的社会心理学研究所里，经常会有关于种族差异和种族不平等的讨论，但是从来没有人提及社会阶层。对，没有就是没有社会阶层之分。当我还是一名研究生理健康的年轻教员时，我在美国参加的大多数会议都会有演讲，演讲内容展示种族间在血压和其他疾病，吸烟、肥胖症和其他健康行为，以及平均寿命上的巨大差异。没有关于社会阶层的。当然，你知道为什么。

到了20世纪90年代早期，我参加的国际会议越来越多。如果发言人来自欧洲，他们的图表总是包含阶层信息而不是种族。美国人就正好相反。然而，到了20世纪90年代中期，事情发生了转变。统计学家开始注意到，在美国，社会阶层的影响度通常高于种族的影响。的确，在过去十几年中，有很多研究已经表明，在吸烟、喝酒、抑郁症、肥胖症，以及所有你可以想象到的生理和心理问题上，社会阶层都有强有力的影响。

社会阶层一般通过人们受教育的年限和他们的年薪来衡量。你可能会认为下层社会的人会存在健康问题，因为比较穷的人没有途径获取医疗保健。事实上，完全相同的社会阶层影响在瑞典和其他国家也存在，而这些国家的医疗保健非常普遍。社会阶层几乎和我们做的所有事情相关，如我们买的食物、我们看的电影、我们选举的人和我们养育儿女的方式，以及可预见的，我们使用功能词的方式。

在过去的40年里，有少数几个研究指出了语言发展、词汇，甚至是家庭语言模式中的社会阶层差异。据我所知，还有以大量成年人为样本来研究语言差异的。近期，我的同事，语言学家大卫·比弗和得州大学奥斯汀分校（UT-奥斯汀）招生办的研究员加里·拉维尼，我们组成了团队来研究高中生为被大学录取而写的入学申请书。UT-奥斯汀是一所择优录取学生的本科学校，每年大概会招到7000名优秀的学生。除了写两篇文章，申请者还必须填写大量的调查问卷才能被考虑录取。除此之外，学生需要提供关于家人社会阶层的信息，包括父母受教育程度和大致收入情况。连续4年分析超过2.5万学生的数据使我们能够研究社会阶层和语言的联系。

下面的表格应该有些熟悉。来自较高社会阶层的学生倾向使用大词、更多的冠词和名词，以及更多的介词。来自较低社会阶层的学生往往更倾向于个人化，在他们的入学申请中使用更多的代词、助动词、现在时态动词和认知机制词汇（很多都与模糊语相关）。

不同社会阶层的大学申请者使用词汇的比例

	中低	中	中高
高社会阶层使用的词汇			
大词	18.7	19.8	20.6
冠词	6.49	6.85	7.05
介词	14.5	14.8	14.9
低社会阶层使用的词汇			
人称代词	11.6	10.9	10.6
第一人称代词单数	8.31	7.98	7.78
非人称代词	5.41	5.03	4.78
助动词	9.18	8.54	8.1
现在时态动词	6.64	5.94	5.56
认知机制词汇	17.9	17.6	17.3
学生人数	4 878	12 885	6 328

据统计，这些影响非常大。这些模式和男女之间、老少之间是相同的。另外，这些都是聪明上进的学生。为什么不同的社会阶层使用的功能词会有不同？显然，学生的成长经历和生活经历影响着他们对词汇的选择。一个很有可能的想法是，从一开始，不同社会阶层的家庭使用语言的方式就不同。一些初步的实验研究支持了这个想法。也许被引用最多和过度解释的一个研究就是贝蒂·哈特和托德·里斯利在20世纪80年代的研究。在研究中，有13个高知家庭，23个工薪阶层的家庭和6个接受福利救济的家庭，他们两年多里每月会有一个小时在家的时间被录音。实验开始时，大多数孩子还不到一岁。我们对录音内容进行了仔细地听写，大人和孩子的所有词都被统计在内。实验设计者发现，最穷家庭中的孩子接触到的词汇数量是高知家庭中孩子的一半还不到。工薪家庭的孩子处在中间。在研究的最后，高知家庭中孩子的词汇量（基于录音）是福利救济家庭中孩子的两倍。表面上看，这是一个具有说服力的研究。然而，即使是实验设计者也同意，解读结果应慎重。比如说，最低社会阶层的家庭只有6个。另外，所有高知家庭的父母中至少有一方是老师，而且几乎所有人都是白领。很有可能，高知家庭的成员对每个月都有学生来家里录音这件事感觉

很自在。而另一方面，福利救济家庭的成员都是非裔美国人，很可能对这个实验的态度和高知家庭成员非常不同。例如，较低阶层家庭的成员可能对实验抱有怀疑态度，所以在录音的时候避免说话。

很可能，较低阶层家庭中长大的孩子和更富裕、受教育程度更高的家庭中长大的孩子语言经历不同。为什么这会导致功能词使用模式的不同，可能与权力和地位问题相关。

谈到性别、年龄和社会阶层：权力之声

在这章中，两组词汇会重复出现。第一个，我们称之为“名词群”，包含冠词、名词、介词和大词。第二个指“代动词群”，包含人称代词（尤其是第一人称单数）、非人称代词、助动词和某些经常和模糊限制短语连接的认知词汇。男性、年长者和较高社会阶层的人使用名词群的频率都很高；女性、年轻者和较低阶层的人使用代动词群的频率都很高。大多数解释这些差异的简单假设是不成立的。例如，某人可能会说，说话更口语化的人可能更倾向使用，比如说，代动词。这对女性和年长者来说可能正确，但不适用于较低阶层的人。也许那些更多使用名词群的人只是读书更多，接触到的词更多。同样，这可能适用于年长者和较高社会阶层的人，但不适用于男性。

最简单的解释是，地位较高、权力较大的人倾向于使用名词群，地位较低、权力较小的人依赖代词和动词。可是，为什么那些拥有更多权威的人使用更多名词相关词汇？还有，同样重要的，为什么权力级别较低的人使用更多的动词和代词？

解决这个问题的一种方法是思考权力和地位能够获取什么。亚当·加林斯基是美国西北大学凯洛格管理学院的一名研究员，他做了一些实验，实验中组内人员会认为自己拥有权力或者没有权力。如果他们相信掌控着自己的命运，就更倾向于自己做决定而忽略别人的想法。那些没有权力的人很容易被人左右。简言之，如果你在某种情况下没有权力，那你最好关注他人。但是如果你是老板，你应该密切关注手头的任务。

正如上一章提到的，人们关注任务时，不会关注自己。事实上，对于大多数任务而言，为了达到其目标，人们必须对任务的主体、进程和具体特征有一个清晰的认识。很像两个人工注射睾丸素很高的人，关注任务的人能够不受人际关系的影响而做决定。

代词和动词的使用不仅代表低地位，还表明说话者更以社交为主。根据定义，大多数代词是社会化的。像“我”“你”“她” 和“他们” 这样的词汇告诉我们，说话者意识到他人的存在以及正在想其他人。第一人称代词单数有些不同，因为它代表着自我关注。在后面的章节中，我们总结了表明人类（和非人类哺乳动物）在其他更有权力的人在场的情况下更多关注自己的证据。为什么动词和特定的助动词也许能够表明地位低，却不是完全清晰。助动词（如是、已经 ）通常是被动表达的一部分：和“他是被球打到的”相比，更主动的说法是“球打到了他”。

不同年龄、不同性别和不同社会阶层使用词汇的相似性只是功能词和社会心理学谜团的很小一部分。一个人是男性还是女性、年轻还是年长、富有还是贫穷，都只是我们身份的很小一部分。语言体现我们是谁，我们才刚了解一点点。在接下来的章节中，我们会挖掘得更深，深入到每个人的内心。

第4章　性格：发现人的内在

使用代词越多的人社交能力越强。缺乏安全感、焦虑或沮丧的人使用“我”字的频率会稍高一些。

你是谁？你，对，你——正在阅读这文字的人——是你的特征、价值观、故事、技能、工作、朋友、人际关系和物质财产，当然，也是使用语言的总和。暂停一分钟，然后思考一下你会如何向陌生人介绍自己。认真点儿。在本章中，你会看到功能词如何揭示出更多你没有猜到的东西。

很多在线交友服务鼓励成员在他们的网页上写一些描述自己的话。碰巧，我和我的学生在我们的研究中已经收集了几千篇这样的自我介绍。下面是一些例子。

我是一个不错的人，非常脚踏实地，不爱闹事，工作努力，关心人，诚实，富有同情心，助人为乐，凭直觉，爱打听，好奇心强。我不肤浅。我不维持肤浅的关系，拥有很好的朋友和美好的生活。

——胡安，27岁的男性

我有一副好身材。我是一个好厨师，有一份很好的工作，房主，特殊的道德感，优秀的决策者，不抽烟，无负担，未婚。我喜欢骑行、皮艇运动、游泳、喷气式滑雪板、登山、滑板滑雪、滑旱冰、骑马。我还很喜欢音乐、电影、话剧、喜剧俱乐部等。

——马库斯，39岁的男性

两个女儿的父亲。我喜欢房子、卡车、狗、孩子；喜欢户外，爱足球和我的女儿们。

——托尼，31岁的男性

***

我是一个浪漫、热情、有激情和喜欢玩的人。对我来说，重要的是家庭、朋友、善良、正直，尤其是笑声。我通过每天的冥想和健康的生活方式来平衡工作和我的精神生活。我庆祝每一天。

——玛格丽特，53岁的女性

我认为自己是一个非常不拘一格的人。我几乎在任何地方都可以有乐

趣，我想尝试所有事情……（也许除了乡村舞蹈）。我的生命中已经有一些独特的经历了，我还会继续发掘生活可提供的东西。我研究古董、针织和钩针，还有我开了13年的汽车，我拒绝换掉它，虽然它的窗户性能不稳定，只能往下摇动，我的刹车灯必须要在电线上缠上胶带才能亮。我是一个化妆师，所以很自然地，我有非常少女的一面。

——琪琪，31岁的女性

跟我熟的人都认为我是一个热心肠的人，是个慷慨的朋友，独立、机智，但是有时候有些固执。

——米拉，34岁的女性

阅读个人广告有点像看电影的某部分。每一个描述都讲述了作者的一个故事，这体现在他说了什么和怎么说的。例如，对比一下前三个男性。27岁的胡安从传统的维度描述了自己性格，这些方面本身是社交属性的：诚实、富有同情心、助人为乐。然后他以拥有很好的朋友这一事实结束。39岁的马库斯描述自己的活动和财产。没有透露他的情感和他人的联系。但是他确实有一副好身材。31岁的托尼描述中有种奇怪的辛酸。在大量词汇中，他表达了一种社交孤立的生活，和他少有的财产、他的狗，此外，最重要的，他的两个女儿。

三个女性的广告也同样说明问题。玛格丽特，53岁的女性，不怎么谈论自己的性格特质，谈了更多指导她的核心价值观。琪琪是一个讲故事的人，你知道吗，她会在你们第一次见面的时候谈论她之前的独特经历。即使是她出故障的汽车窗户这个古怪的故事也是她的一部分。最后一个女性米拉，也许她都不相信自己对自己的看法，所以她告诉读者别人是如何评价她性格的。虽然所有这些人都是有意识地在社交网站上宣传自己，但是我们还是能够看出一些他们的特点。

除了显示一些每个人的身份特征外，不同的广告指出人们关于性格思考惊人的不同。有些人的核心是他们的财产、工作、技能、特征，或者主导特征、他们的价值观，抑或他们的故事。

如果回过头看这六个人的广告，你会发现每个人使用功能词的方式大相径庭。普通的性别差异出现了：女性使用更多的代词（尤其是“我”字）、助动词和认知词汇，男性使用更多的冠词和名词。但是即使是男性和男性的描述，或者女性和女性的描述之间差异也很大。对于几乎任何一类功能词或情感词汇，每个人都有一套独特的使用模式。正是这些不同的模式为他们的世界提供了线索。

很有可能，那些贴出广告的人去约会了。他们可能分享了更多自己的事情，作为回报，他们也期待约会中透漏更多自己的信息。在约会过程中，他们肯定会谈论自己的日常生活、最近的天气、食物、衣服，或者其他吸引他们注意力的话题。同样的讨论在面试中、第一次见办公室同事，或者只是在聚会上和人聊天时也会发生。我们一直在直接和间接地分享和寻找关于性格的信息。那么有人可能会想，现代心理学应该用一些快而有效的方式对人们讲述的自己的故事进行分类。这么想就错了。

通过倾听人们说的话来研究人的性格是极其困难的。每个人故事的内容和类型都不同。在线交友中，有人写了好几页，有人只写了一两行。一个人只是列出了自己的工作和财产，另一个人列出了自己的生活目标，你如何对比这两个人？如何对人的自我描述进行整理和分类？在过去，研究员会聘请一些评价者来阅读每一篇自我评价，然后对文章作者的性格、写作风格、目标和兴趣做出判断。这个过程缓慢、靠不住且非常昂贵。

最近，我的同事们和其他人已经开始应用一些有前景的计算机方法来分析开放式的性格描述。正如你可能会猜到的，功能词的不同模式揭示了人的性格和思考方式的重要部分。

通过功能词了解人的性格

20世纪80年代早期，一名年轻的南加州大学语言学系的学生道格拉斯·比伯决定研究就语言使用而言，不同的文学体裁有何不同。

例如，文学小说、非小说类书籍、戏剧、神话，甚至是爱情小说，它们的作者使用的文字、语法和句法有不同吗？道格拉斯·比伯偶然发现一种不常用的统计学方法，称为因子分析（稍后进行更详细的介绍），他最终出版了名为《口语和书面语的语体差异》（Variation Across Speech and Writing）的书，书中回答了这个问题。当然，英语学者很早就能够判断神话和爱情小说间的区别—情节不同。但是比伯的方法基于计算机，他通过代词使用量发现，爱情小说更为个人化，使用的现在时态动词比神话和其他任何体裁都多。事实上，他发现几乎每一种体裁都有自己独特的语言风格。

20世纪90年代中期，我无意中读到比伯的书，被他的想法所吸引。虽然他关注的是如何通过分析部分语言来区分伟大和不那么伟大的文学作品，但是没有理由不能以同样的方式通过观察人们的写作来区分他们。这次，我和我的一个前同事合作，劳拉·金来自密苏里大学，是世界上最富创造力的人格心理学家之一。比伯书中分析词汇是为了区分文学体裁，而我和劳拉在想，我们能否通过人们每天的写作来区分人的风格。也就是说，不同的写作风格能反映不同的性格吗？

第一步就是要获得上百甚至上千的写作样本，写作内容是一般话题。幸好，我照例每年都会教普通心理学的大课。每个学期开始时，我会向学生介绍意识和思维的本质。1890年，现代心理学先锋之一威廉·詹姆斯写了大量关于思想本质的文章，把它比作溪流。一种想法或感觉会自然地引发另一种想法或感觉，然后会引发另一种想法或感觉。每一种想法和感觉激发下一个，下一刻的想法和上一刻的想法可能完全无关。进一步讲，如詹姆斯所说，我们的思维沿着意识流动，我们只能在某一刻意识到一个单一的想法。

詹姆斯之后，有很多研究者进行意识流的实验，实验中让人们谈论或记述脑中不断出现的想法。多年以来，我一直要求我的学生在写作时跟踪他们的思路20分钟，为的是让他们更好地了解威廉·詹姆斯和意识流本质。虽然这个练习要求学生在线提交，但是他们的写作不会被评分。我们现在有接近8000篇文章，并且已经得到学生们的许可，可以用来做研究。

为理解意识流写作，想象一下如果你在跟踪自己的思绪和感受时会如

何写作。这个指令非常简单，就是不停地写下你脑中的想法和你的感受。不要停，就是写。下面是两个非常标准的例子，出自两个大学生之手。

冰箱正在发出奇怪的声音。穿过走廊的人们说话声音好大。有些人放音乐的声音好大。我的胃在咕咕叫。我想，我想要一份快餐。我头疼。我的嘴好干。我想喝点东西。我屋里需要更多的瓶装水。我需要重新涂我的指甲。颜色都掉了。我手上的戒指太大了，需要调整一下大小。

另一个：

我感觉不太舒服。不知道尼克是否安全到家。也许我会打电话看他是不是安全到家了。想到心理测试，有些紧张。我本来应该今晚看书的，但是我生病了。我想家。我想我的家人和我的狗狗们。巴兹很可爱，雷德总是叫，但是我依然爱它。我希望周四考试的时候病已经好了。

在这两个例子中，人们都是想到哪写到哪。也就是说，他们在想到这些的那一刻就记录了下来，没有经过构思和分析。我们通常认为这种类型的思考是“即时性”的一种。偶尔，我们所有人都会这样思考。有些时候我们甚至会这样说话（也许在我们累的时候或者在跟密友闲聊的时候），使用小词、现在时态和高频率的人称代词，尤其是“我”字。

我们从数百篇文章中分离出各类功能词，包括人称代词、冠词、介词、连词和助动词。然后我们用称为因子分析的统计学方法来观察出现了哪些功能词块。意识流文章中出现了三个表明不同写作风格的因子：正式性、分析性和叙述性。事实上，不论分析什么文本，几乎在任何体裁的写作中我们都发现了同样的因子，包括同类型文学作品、歌词、入学申请书或是绝命书。

◎正式性

正式性是常出现的功能词因子，正式性通常表现为呆板，有时表现为没有幽默感，还有一丝傲慢。可以认为正式性与即时性相反。高度正式（或低即时性）的思考或写作通常包含大词和高频率的冠词、名词、数字和介词。与此同时，正式性写作很少使用“我”、动词（尤其是现在时态动词）、差异词（例如会、应该、可以）和常见副词（确实、非常、所以）。下例是被计算机识别为高度正式的文本。

我听到凉鞋摩擦的声音。电视和广播模糊的声音融合成几乎不连贯的杂声和节拍。一段独白出现在客厅的电视上。沐浴在来自客厅的众声音之中，嗡嗡声——TruAir空气净化器的声音吸引了我的注意力……宇宙谜团出现在我的脑海中，为什么会有热寂说？

与即时性例子相比，高正式性作者更有文化，更有距离感，给人一种很严肃的感觉。有趣的是，比伯在做体裁分析时，提出的第一个因子也是正式性。浪漫小说即时性很高，与其相比，学术写作和非小说类作品的正式性更高。

说话或写作的正式性与一系列重要问题相关。正式性更高的作者更关心地位和权力，较少进行自我反思。与没有那么正式的作者相比，正式性更高的作者较少抽烟和喝酒，心理更健康，但是对自己和他人的诚实度较低。随着年纪的增长，人们的写作和说话方式从更直接转变为更正式。换句话说，功能词的第一个因子具有非常大的社会暗示和心理暗示。

◎分析性

分析性因子用来识别谁在努力理解他们的世界。分析的标志就是做区分。这些区分可以是人们做什么和不做什么，他们通过了测试的哪部分和没通过哪部分。有助于分析性思维的词包括排除词（除了……、没有……、除……以外）、否定词（不、从未）、因果词（因为、原因、效果）、洞察词（认识到、知道）、假设词（也许、可能）、确定词（绝对、总是）和限定词（一些、很多、更好的）。分析性思考的例子如下所示。

我一直都知道每个人都是不同的，但是我想我是因为和不经常一起玩的朋友在一起才会想到这个问题。很奇怪，有些事对某个人来说非常普通和简单，但对另一个人来说就非常陌生和复杂。谁知道我为什么会有现在的思考方式，但是看来这样的思考方式和我遇到的大多数人不同。有些人非常需要被关注，他们会做一切来吸引别人的注意力。我不确定他们是否应该被责备，因为这也许是他们知道的唯一表现方式。

写这段话的女性正在试图解析这个世界。虽然她不是一个优秀的作者，但她试图理解是什么造成了人与人之间的不同。分析性因子反映了认知复杂性。说话和写作时做区分的人在大学的成绩会更高，他们可能更诚实，对新事物的态度更开放。他们阅读更多，与分析性弱的人相比，他们对自己的认识更为复杂。

◎叙述性

有些人天生就是讲故事的人。他们控制不了自己。从一个简单的语言视角看，故事讲述涉及的功能词包括人（意思是所有类型的人称代词的使用，尤其是第三人称代词）、过去时态动词和连词（尤其是包含词，如随着、和、一起）。下面是一个有趣的例子。

好的，我的朋友克莉丝这周末要来市里看足球赛。她想拥有一段美好的时光，所以我们周五晚上出去玩，而她简直全浪费了……她吐遍了我们去的任何一个聚会和每一个厕所。我们被松饼屋的人赶了出来……赶出来！说真的，谁会从松饼屋被人赶出来……真是疯狂。

被人从松饼屋赶出来？确实疯狂。虽然写作任务是让人追踪自己的思绪和感受，但是至少有20%的人忍不住去讲述某种类型的故事。事实上，叙述性因子得分高的人更懂得社交，朋友更多，且他们评价自己是外向的。

简单考虑一下这些结论意味着什么。用数据统计方式将人们使用的功能词进行分类，通过分类我们观察到他们的思考方式、组织世界的方式和与他人连接的方式。几乎不被人发现的功能词揭示着人们的本质。

随着劳拉和我研究人们思考方式的深入，我们发现人们的思考方式和写作方式惊人的一致。例如，学生意识流写作练习的写作方式和他们关于神经系统运作论文的写作方式相关联。其他研究表明，你年轻时的写作方式和后来的写作方式保持一致。如果你曾经无意中看到了自己年轻时候写的作文或日记，你会发现自己的写作和思考方式没有太大变化。换句话说，人们的语言方式是人的一部分。

人们的思考方式会发生变化，取决于他们和谁在一起、做什么和对自己的感受。在后面的章节我们会讨论到，一个处在沮丧情绪中的人的正式性会发生从高到低的变化。当一个人尝试做出他（她）一生中一个重要的决定时，他（她）的邮件、博客或自然会话中会显示出分析性增高的迹象。

总之，我们可以通过一个人的写作或说话方式研究他，并且更多地了解他。举个例子，回过头来看一下本章开头几个人的个人广告。带着你对思考方式的相关知识，你可以确定谁的正式性最高（马库斯和米拉），谁的分析性最高（胡安和米拉），以及谁的叙述性最强（玛格丽特和琪琪）。这些分析都是基于少量的单词，所以在解释时应谨慎；收集的文字量越大，结论可信度越高。当我们回到爱和关系的研究时（第8章），你会看到我们是如何匹配思考方式的，并且能够根据双方的写作和说话方式来推断哪些关系可能更持久。

工具栏：因子分析的魔力

有些阅读本书的人不是很喜欢图片或表格，当然更不喜欢统计。即使你打心底讨厌统计，你也许也会发现这个工具栏里的内容很有意思，非常吸引人，真的。

因子分析基于相关性。一个简单的相关性是一个统计数值，它可以告诉我们任何两个变量是如何彼此相关或共变的。人越紧张，通常他们的内疚感越深。紧张程度和内疚程度通常是共变的关系。我们知道这个关系是通过调查问卷测试人的紧张程度和内疚程度，然后将两个选项关联。所得统计数值，即相关系数，从数据角度告诉我们这两个概念是如何紧密相连的。

因子分析是更精细的相关方法。它不是使两个变量关联，而是变量组之间的相关。所以，想象一下我们的问卷是让人们评价他们的紧张程度、内疚程度、害怕程度、悲伤程度、热情程度、积极程度和高兴程度，我们可能会找出两组相关群。紧张程度、内疚程度、害怕程度、悲伤程度应该彼此紧密相关。一个人紧张的时候很可能感到内疚、害怕和悲伤。第二组应该包括热情程度、积极程度和高兴程度。因子分析决定问卷最初的7项内容如何自然地被分成更大的因子。在这个例子中，因子分析会找出两种截然不同的变量群：消极情绪和积极情绪。

因子分析很奇妙，因为它可以帮助研究人员把一大群变量浓缩为易于管理的重要因子。例如，比伯的语言分析依靠语言学家关注的几十个语言维度——代词、名词、短语、分词、物主代词。通过因子分析方法分析400多本书和文章中使用的语言维度，他把几十个语言维度缩减整合为7个维度。他的相对简单的因子能够用来区分文学体裁。正如下文所描述的，因子分析也是研究性格的重要方法。

通过图片发现性格

我的同事山姆·高斯林是一名富有创造性的研究员，他在研究“行为痕迹”。他尝试通过观察人们的办公室、卧室、网页、书籍和音乐收藏来猜想人们的性格。他发现，无论人走到哪里，都会留下性格的痕迹。死板的、谨慎的人通常办公室和卧室都非常整洁，音乐收藏排放整齐。这样的人甚至会有一个详细的整理邮件的方法。

我们的所说所写同样可以被认为是行为痕迹的一种形式。功能词是很好的指示词，显示出人们与他人连接的方式、思考世界的方式，甚至是思考自己的方式。同样，内容词也可以产生有价值的线索。为什么你的一个朋友一直在谈论他的猫咪，而另一个朋友感觉有必要谈谈她的饮食？谈话内容体现人们认为重要的问题，包括他们的价值观、目标和性格。

谈话内容体现人们关注的东西。几年前，那时候还没有研究语言，我对人们是如何看待他们的世界很感兴趣。人们走在路上时，他们关注周围环境的方式是否相同？我们对同样物体和事件的看法有多大差异？在一个小实验中，我在棒球帽上安装了一个摄像头。我的几个学生是我的小白鼠。实验中，给所有人的指示如下。

戴上你的相机帽以后，从走廊尽头的楼梯下楼，然后离开大楼。走到学校附近两条街外的购物区。走进拐角处的杂货店，然后买一包口香糖。离开商店，然后走心理学楼后面的小道返回实验室。就像平时一样，忘记你的帽子。

这一段路对于5个学生和我来说大概花费10分钟。每个人都按照指示走了一遍，然而视频中的景象却完全不同。看完视频后，所有人对每个人看周围的方式不同而感到惊奇。一个不太自信的学生大部分时间都盯着地面，几乎不抬头看路人的脸。两个年轻的学生会仔细观察每个异性的长相。另一个学生在口香糖面前停滞了一会儿，对比不同牌子的口香糖价格。有一个视频让我印象尤其深刻，携带它的人是我们里面非常高的一个人，他比我高出30多厘米。他扫过了每个人的头顶，这个角度是我从来没有想象过的。

我的大多数学生会猜测哪个视频出自哪个人，因为他们了解每个人的兴趣、观念、性格和高度。每个人看物体、人和世界的方式都不同。他们处理周围信息的方式不同，因为每个人在走路过程中吸收的信息不同。如果我仔细对每个学生进行采访，他们会用不同的词汇描述他们到杂货店的这段路程。我想，如果分析他们的内容词，我们会发现很多关于他们的信

息，正如我们根据他们使用的功能词发现了很多信息一样。

◎透过一个瓶子看自己

相机帽实验几年后，我和我的一个研究生学生聊天，她叫辛迪·钟。辛迪是加拿大人，之前生活在多伦多。她最近刚搬到得克萨斯，她发现这里和她之前所到或所想象的任何地方都有一些不同。可能是因为那时得克萨斯天气炎热，所以她走到哪里都带着一瓶水。我想她是觉得沙尘暴随时都会来临，所以有必要做好准备。当我们聊到语言和感知的本质时，我们俩都提到了人们对她的水瓶可能会有不同的看法。不知不觉，她的水瓶成了我们谈话的焦点。实际上，我们拍了一张水瓶的照片，还让别人来描述它。



辛迪的水瓶

自从2002年那次谈话后，上千人看过这个瓶子并且参与了5分钟简短介绍瓶子的活动。如果你想要试一下瓶子实验，请访问www.TheSecretLifeOfPronouns.com ，然后点击“Perceptual Style：You are what you see”。实际上，我鼓励你在进一步阅读之前做一下这个练习。我们等你。如果你没有访问网页，下面是练习的指示：给你5分钟时间，请向一个从未见过该图的人描述一下你所看到的画面，要不间断地写。图片下面是写作区，写作区上方有一个小的计时器，告诉你5分钟时间已经过去了多少。看一下瓶子的图片然后想一想在接下来的5分钟里你会写些什么。

人们描述瓶子的方式差别很大。以下是一些节选。

这是一个透明塑料瓶，外观类似一个水瓶，它有一个白色盖子，瓶身上有一个红色的标签，我可以看到前两个字母和后两个字母，应该是OZ和KA。红色标签上的字母上面有一幅图。上面还有一个蓝框，里面的字很小几乎认不出来。成分和公司信息是黄色的小字……

——61岁的女性

这是一个小塑料水瓶，上面有红色的标志。它的顶部是白色的。瓶里还有一些水。大概有三分之一的水。光从右面来，所以瓶子的影子落在瓶子的左面。部分瓶身有反光，因为有光……

——25岁的女性

这是一个干净、无色的水瓶。大概20厘米高，很小，你的手可以握住它，而且还能余下一个手指。它顶上有个拧上去的盖子，可以保证在打开瓶子前里面的水是新的，而且可以防止水流出来，所以你可以在喝完水后盖上盖子……它是空的。里面没有水，所以我希望你不渴。不过这使得它很轻，所以很方便携带。当它是空的时候，也很容易打翻它。

——60岁的男性

每一个人都强调瓶子的不同面。第一个人几乎只关注了瓶子上的字和字母。第二人描述了瓶子投射的影子。第三个人传达了握住瓶子和用它喝水的感觉，口渴的感觉和它的重量。

这个瓶子项目影响了我关于性格的思考。随着描述瓶子的人越来越多，辛迪和我开始用新的方法分析他们的描述。开始的时候，我们把功能词和了解到的人物性格联系起来。比如说，使用代词（如“我希望你不渴”）越多的人社交能力越强。缺乏安全感、焦虑或沮丧的人使用“我”字的频率会稍高一些。显示出正式性的人在他们的描述中高频率地使用冠词和介词，他们常常年纪更大，做事更有条理、更谨慎。

但是这些分析落下了某些东西。回头看下这几个例子。是的，这三个人使用功能词的方式不同，但是更惊人的是，他们的关注点不同。为什么一个人关注的是颜色，而另一个人想的是瓶子拿在手里的感觉？人们在自然状态下描述一个物体的方式体现了他的思考和感知方式。例如，我自己在完成这项练习时，从没考虑过去描述瓶子拿在手里的感觉。

我们需要一个更有效的方式来分析人们是如何观察和感受这幅图的。几周以后，辛迪在大厅叫住了我，她想到一个办法。这个办法简单而巧

妙。这项技术之后被称为意义提取方法，基于我们收到的所有描述瓶子的文章中常用的内容词。

为了理解这项技术是如何实施的，请再次看一下上面三个描述瓶子的文本。正如你能想到的，常用的功能词有：这、它、是、一个、的。另外，有很少一部分内容词：水、瓶子、红色、小、白色、盖子、干净、容易空、标签、左面、字母、轻 和塑料 。如果说人们关于瓶子的文章大概有1500篇，那么会有更多的内容词。事实上，我们发现，1500篇文章大概有2万多词，而大概只有175个内容词比较常用。

现在想象一下，我们深入每一篇文章，然后观察每个人是如何使用这175个常用词的。接下来，我们观察这些内容词是如何自然地聚在一起的。如果你觉得这听起来怎么像因子分析，那你就对了。因子分析可以用数学方法确定不同作者在关于瓶子的写作中会同时使用哪些内容词。我们发现，如果一个人在文中使用了“黄色” 这个词，那么这个人倾向于使用其他的颜色词，如“蓝色”“绿色”“白色” ，也倾向于使用“天空”“山” 和“标签” 。当你看到这些词时，会立刻意识到这些词是人们描述瓶身上的标志时常用的词。

另一个词群包含“圆柱体”“形状”“圆锥体”“顶部”“高” 和“宽” 这样的词，以及其他和瓶子形状相关的词汇。而另一个突然出现的主题包含“光”“灰色”“影子”“反射”“背景”“墙” 和“桌子” 这样的词：背景/照明主题。甚至还有一个主题，这个主题中人们沉迷于盖子是否已开封，以及握住瓶子或用它喝水的感觉——有点儿像上面的第三个文章。

如果你不是计算机和因子分析方面的专家，你可能会想：“呃！为什么你需要一个计算机程序来判断一个人在谈论颜色还是形状？这很明显啊。”相信我，对于计算机来说这不明显。意义提取方法非常神奇，它可以从大量相同的文本文件中自动提取出所有主题。计算机（而不是人）可以识别人们写作中暗含的主题。

从心理学角度来看，它能让我们看出不同类型的人在写作时突出哪些主题。即使是从最浅显的层面看，结果告诉我们有些人关注颜色，而有些人不关注。有一组人描写了字和字母，而其他人没有。

我们还深入研究了描写瓶子不同主题的人的行为。最有意思的是那些描写照明和投射在瓶子左面的影子的人。大学生中，那些描写影子的人考虑更周全，更有艺术气质，不太在意外观。他们成绩更好，参加更多的艺术演出，电脑游戏也玩得更多。还有，他们不太可能饮酒过度，会做任务清单、打扫卫生、修东西、泡澡（尽管他们冲澡的次数和其他人一样频繁）、吹头发。

还出现了其他模式。那些关注标签上字的人更可能是女性，更可能看书较多。那些对瓶子的材质和感觉感兴趣的人存在各种类型的身体症状。那些对瓶子的外观进行描述的学生似乎在自己和他人之间设立了一道屏障。他们不怎么合群，很少与人交心，而且如果经历了大的波折，他们会选择不和别人交谈这些事。

谁会想到，描述一个水瓶这种无关紧要的事情也可以揭示如此之多。虽然很有前景，但是结果模式不是非常有力，尽管它们的数据非常可靠。换句话说，如果你要求一个朋友描述这个瓶子，当他提到图中的影子后你就去买了吹风机给他，这会非常鲁莽。相反，把这些结论当作是我们分析语言的线索。想象一下，如果我们要求人们描述更为复杂的事物，我们会看到什么。

◎透过他人看自己

我喜欢这个瓶子测试。受到这些结论的鼓励，我和我的学生尝试了另外一张图片。这次我们想要一个更复杂、社交性更强的刺激。我们需要一张普通的真人照。为了找这样一张图片，我翻遍了我的抽屉。突然间，我发现两个非常好的朋友在户外派对上的照片，这张照片我和我妻子保留好几年了。经过他们的同意后，我开始在几次研究中使用这张照片。



在进入绿色花园项目之前，请先看一下这张照片。花几分钟时间描述

一下这张照片，写给另外一个没有见过它的人。如果你决定不写下自己的描述，至少在脑海中想一想，并记住你都关注了什么。

相比其他人，这张图格外吸引我的学生凯特·尼德霍夫。凯特的研究方向横跨传统社会心理和市场营销。她看完照片后想，人们描述照片的方式是否能够预示他们的购买模式。更广泛地说，她想知道人们的日常行为是否能够反映他们看待世界的方式。凯特最终创办了自己的营销公司，她建立了一个网址，来自各行各业的数百人在上面描述这张照片，并且完成了一份关于自己日常行为的详尽调查问卷。（一个简短版的调查问卷可以在www.TheSecretLifeOfPronouns/LIFE 填写。）

很多人都对照片进行了描述。类型千差万别。下面是一些例子。

这是这个女性的房子。她婚姻幸福。这个男性正在给她讲故事或者黄段子，因为这个女性很明显对他所说的东西表示很惊奇。她喝的是自制的红茶。这两个人不是夫妻，但是各自都已婚。

——22岁的女性，最近刚订婚

这位女士刚从屋里出来，她从刚打开的酒桶里给自己倒了些酒。她已经喝了很多，动作看起来有些做作。这个男性看上去对抢走她手中的啤酒更感兴趣，而不是跟她接触。他想夺过她手中的酒杯然后离开。

——19岁的男性，饮酒过度。

有两个人在房子的后院聊天，这里正在举办烧烤派对。她穿着一件长袖、牛仔布、带扣子的T恤（估计在GAP买的），塞在裤子里，袖子卷了两下到小臂位置。她的头发是棕色的，剪的很有品位（有层次），中等长度，到脸颊位置。她还带了一条黑色腰带（古驰？），右手腕上戴着一个红色手镯。

——25岁的女性，每月花3000多块在衣服上

图中的两个人正在深入地讨论政治。这个女性得知男性是共和党，正在表示惊讶。她对他辩论的逻辑感到震惊……这个男性正在表达（用双手）他欣赏“更小的政府”，他认为这归因于共和党。

——21岁的女性，活跃在政坛

意义提取工具从人们的写作中提取出一些主题。常见的主题包含：两个人的社会联系（他们在约会、已婚，他刚说了什么），对衣着的分析（从简单的描述到时尚角度的观察），佩戴物（他们的表、他的眼睛、她

的唇膏），这个女性的饮品（酒、茶，她喝多了？），后院（花、树、蹦床），房子（床、壁板、屋顶），描述这两个人（她的面部表情、他们的种族和年龄）。和瓶子实验一样，每一个主题都有一群相关词汇。

正如你看到的，人们描述的主题实际上与他们自己的生活相关。正处在恋爱关系中的人们，尤其是女性，写作内容是有关这两个人可能关系的，且通常会对他们戴的婚戒进行评论。那些在自己的衣服和服装配饰上花销很大的人，更可能描述这两个人的衣着和配饰。园丁评论花和树。喝酒越多的人，越可能提到这个女性正在喝酒并且有些醉意。

词汇使用可以预示购买行为这件事对谷歌或者其他搜索引擎公司来说并不新鲜。如果你使用谷歌邮箱之类的产品，然后收到一封来自朋友的邮件，邮件内容是他早餐要吃葡萄干麦片，那么你可能会发现，你的邮件界面出现了一些关于果干、芒果冰或减少胆固醇的药的小广告。今天我收到家人的一封邮件，他一个朋友40岁了，最近有些不理性和易怒，他问我该怎么办。蹦出来的广告有“治疗青春痘”“青少年最佳读物”和“再要个宝宝吗”。显然，统计学上有一些词（也许像“不理性”和“易怒”这样的词）与青少年和孕妇相关联。

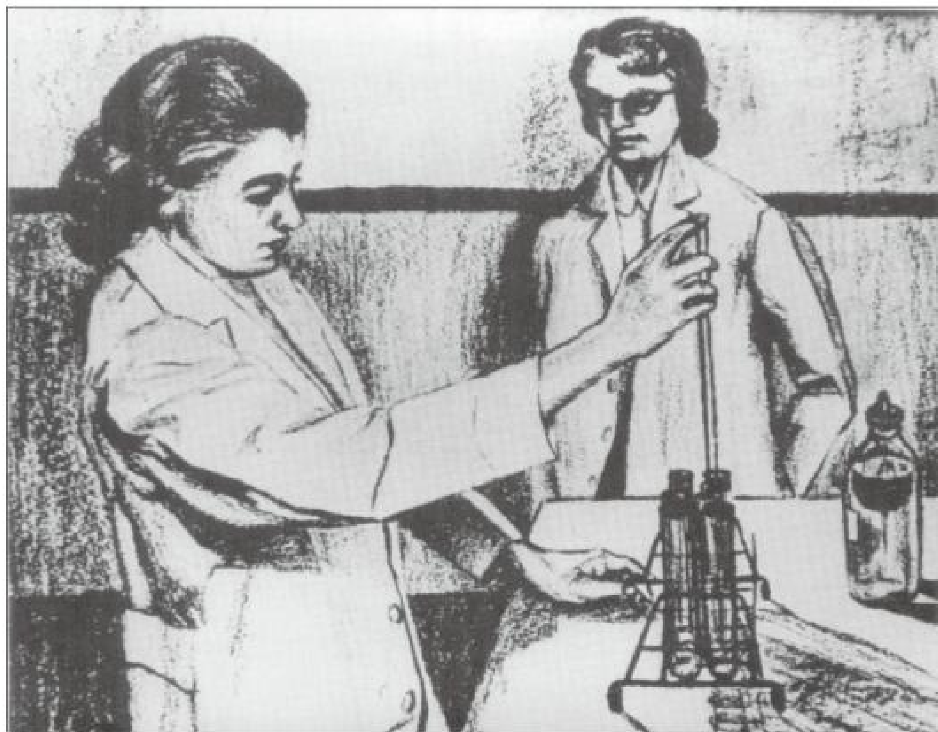
◎投射测验

瓶子和花园派对的图片测试是一个古老心理学想法的变体。西格蒙德·弗洛伊德、卡尔·荣格，尤其是弗洛伊德的女儿安娜·弗洛伊德声称，自然状态下，人们会将自己的想法和感受投射在别人或别的物体上。一个人可能因自己错过一场约会而生气，他遇到一个老朋友，然后会说：“发生了什么？你看上去有些生气。”通过投射防御机制，人们可能无法感知自己的愤怒，但是他们会在其他地方看到愤怒。

20世纪20年代早期，年轻的德国心理学家赫尔曼·罗夏进一步论证了这个想法。他开发了一个方法，让人们观察墨迹的模糊图片，然后描述他们看到的东西。罗夏测试基于的想法是，人们内心深处的情感和担忧会投射在墨迹上。例如，一个人长期生活在有暴力倾向的父亲（或母亲）的恐惧中，他可能会从一个墨点中看到一只凶猛的熊，而另一个生活在幸福家庭中的人可能会从同一个墨点中看到一只蝴蝶。继罗夏早期成果之后，成千上万的临床医学家依靠罗夏测试来探索人们的心理状态。有很多研究者质疑罗夏测试，因为人们的反应是不靠谱的，而且会根据测试实施者的不同产生很大变化。我不禁想，这个方法是否可以通过计算机得到改良，利用计算机执行测试并为测试打分。

20世纪30年代，哈佛大学的亨利·默里和克里斯蒂安娜·摩根开发了另一个投射测试——主题统觉测验，或称TAT。这个测试不是让人说出他们

看到墨迹后想到的第一个事物，TAT鼓励被试者根据一系列模糊的图画来编故事。据说看图编故事可以反映出被试者潜在的心理问题。下图是标准的TAT刺激图。通常会要求人们看图片几秒钟，然后描述他们认为图中发生了什么事情。



你也试一下。如果你想在讨论之前做一下这个测试，请访问：www.TheSecretLifeOfPronouns/TAT 。在线版本中，你会收到电脑处理的反馈，这基于我们开发的试验性方法。即使不访问上面的网址，你也可以看这幅图片，然后简单编一个故事，描述一下图中正在发生什么事情。每个人都在做什么、想什么？过去发生了什么，以及未来可能发生什么？

TAT的一个用途是检测基本的心理需求。受默里的影响，戴维·麦克利兰开发了一个模型，假设我们所有人都受三种基本需求所驱动：成就、权力和归属。通过不同的TAT图片，麦克利兰确定了每种需求在影响行为上的优势。你刚才看到的那幅图和三种需求相关，但最初是用来评估成就和权力需求的。TAT线上测试已经吸引了来自世界各地成千上万的人。下面是针对图片的两个回答。

这个故事发生在雪莉和她的新员工之间，雪莉是一个保守的检查员，她的新员工刚到公司不久，她们之间起了一些冲突。索尼娅投入工作的热情度不高，雪莉对此感到不满意，准备让索尼亚难堪，作为一个检查员，她很不喜欢这种情况……雪莉的身体一点儿都不放松，因为她的注意力没有集中在实验上，而是集中在她准备轻视的这个人身上，因为这个人威胁到了她不可动摇的地位。索尼娅成功地完成了她的任务……且作为奖励，她得到了新部门实验研究员的职位……雪莉从未被认为是……忠诚可靠的人，虽然这些年来她一直试图向组织证明。

——64岁的男性

朱莉早上起来，想到今天要去实验室，又会是艰难的一天。她的母亲负责看管她，母亲非常严格、守旧，认为朱莉的一切想法都是耽误时间……化学物质混合几分钟后……烧杯融化了，化学物质开始冒泡沫，泡沫弄得哪儿都是，毁掉了所有东西。她的母亲大声责骂她……但是这次朱莉受够了她的母亲……她终于决定告诉母亲这些年来自己的感受。奇怪的是，母亲听了进去……她们两个都松了口气，终于把事情说了出来，并且有机会重新构建良好关系。

——17岁的女性

这些故事和大多数的TAT故事都有一定的透明度。第一个故事表明，有人在工作中会感觉受到了年轻人的威胁，会有一种无力感。按照麦克利兰的理论，这个故事暗示被抑制的权力需求。事实上，根据麦克利兰理论，一种检测被抑制权力需求的方法就是看一个人使用否定词的频率，如“不”“从未”。这位64岁的绅士写这篇文章时使用了五次否定——令人印象深刻。有没有可能这个故事能反映出他生活中一些类似的经历呢？

同样，你觉得写第二个故事的17岁女孩有没有可能正和自己的妈妈有些矛盾呢？当然，只是猜测。第二个故事显示，作者高度需要归属，中度需要权力和成就。

对于了解人们需求的评分方法有点儿复杂。以往，这些文章由经过专业训练的评分员一字一句地阅读并打分。最近，计算机程序被开发用来系统地寻找表示成就（“赢”“输”“成功”“失败”“努力”）、权力（“威胁”“老板”“员工”“领导”“跟随”“控制”“服从”）和归属（“爱”“朋友”“孤单”）的词汇。需要注意的是，意义相反的词可以表示同样的需求，如“输”和“赢”。执着于成就的人会渴望成功，同样，也会害怕失败。对成就需求低的人基本不考虑与成败相关的事情。

关于成就、归属和权力需求的研究得到了重要的发现。例如，那些抑

制权力需求的人血压较高。大卫·温特是需求分析的领先者之一，他分析了世界领导人的演讲，能精准地预测领导的风格、宣战的可能以及其他行为。例如，约翰·肯尼迪和乔治W.布什第一次就职演讲显示出两个人对权力和归属的高度需求。在大卫·温特看来，一个有权力的领导在做重大决定时依赖一群紧密联系的好友是一个有毒的组合。2001年年初，分析完布什的第一次就职演讲后，温特警告说，布什的语言具有攻击性，和他的一群反抗反对意见的紧密追随者的基调一致。

◎实验室内外推断性格

像是瓶子测试、花园派对测试、罗夏测试和TAT这样的方法都旨在通过研究人们使用语言的方式来推断其行为或性格。这些方法通常都在可控的实验背景下完成。人们看着同样的照片，遵循同样的指示，被要求以同样的方式记述或描述它们。当每个人面临的情境几乎相同时，任何语言上的差异都应该反映人们本身的差异。

对研究人员来说令人沮丧的一件事是，在人为限制的实验室之外无法解读这些结果。例如，在瓶子测验中，那些描述瓶子投影的人更有艺术气质，是更优秀的心理学学生，对干净和整洁没有那么执着。这是如何解读的？很有可能，我们偶然发现了一种感知方式，这种方式只有在人们解读某些视觉场景时才会显现。换句话说，瓶子测试结果告诉我们，有些人如何天生就能从物体的视觉特性中察觉出细微差别。有可能是在那张特定图片的特定灯光下，那个特定的瓶子使一些人注意到了它的投影，使另一些人注意到了其他东西。

说得更宽泛些，我们只通过人们日常的语言就可以得到同样的信息吗？这可能包括他们的邮件、对话、微博或专业写作。答案是肯定的。这就是本书的核心内容。分析型和分类型的思考者在描述一个无聊的瓶子、讨论后院派对时，或者和一个邻居聊吉利威蒂先生的胃病时，更倾向于使用冠词、介词和否定词。当然，我们说话和思维的改变取决于我们所在的环境。在正式场合，我们都讲话更正式；在疯狂的派对上，我们讲话会，嗯，很疯狂。即便如此，我们不论走到哪里，不论处在什么情境，都会带着自己的性格，都会在功能词中留下痕迹。

通过语言了解他人和我们自己

通过这章的结论，有没有可能对你来说可以更好地“阅读”他人？语言的研究工作能帮助你成为一个更好或效率更高的人吗？让我们对这两个问题肯定地回答“是的”。

简略的勾画：奥萨马·本·拉登的语言

考虑一个公众人物的语言，比如本·拉登。他的大部分成年生活中留下了许多采访和演讲的记录，以及书信和文章。不论他的语言是阿拉伯语还是英语译文，都显示出高度的自信，甚至是自大（“我”字频率非常低，“我们”和“你们”的频率很高）。本·拉登不像其他阿拉伯极端组织领导人，包括他的好朋友艾曼·扎瓦赫里，本·拉登是一个故事讲述者（叙事标志较多——过去式动词、社交指代），带有一股严厉和敌对的锐气。我们整个分析将他视为一个高度需要权力，中度需要成就，低度需要归属的人。辛迪·钟的意义提取方法显示，他生命中真正痴迷的事情从对家乡沙特阿拉伯的愤怒转变为美国对伊拉克和阿富汗的入侵。有趣的是，与基地组织的同伙比，他对以色列的兴趣非常低。至于他是否喜欢黄昏在海滩长时间散步，没有数据显示。

◎通过词汇使用阅读他人

还记得第2章议员约翰·克里的故事吗？他的助手鼓励他在演讲中使用更多的“我们”，少用“我”。他的智囊团错误地认为一个人在演讲中使用“我们”会拉近说话者与听众的距离。这对任何想要通过分析语言了解他人性格的人来说，都应该把它当作一节示范课。词汇计数的首要原则：不要相信你的知觉 。

如果你想通过别人的语言来了解他，那么你需要真正地去数词汇。你可以手动做这些，但这会是一个缓慢而痛苦的过程。或者，你可以使用计算机程序。当然，简单的计数只是解码他们性格的第一步。

你想了解他人的哪方面？正如本章概括的，检测一个人的思考风格是一件相对容易的事情——无非就是正式型、分析型或者叙述型。后面的章节为检测欺骗、领导力、社交能力等提供了线索。难的是通过技术（例如意义提取方法）来了解什么吸引了人们的注意。幸运的是，与功能词相比，我们的大脑更擅长捕捉内容词，所以实际需要计数的词汇会少一些。

如果你想知道什么主题引导着人们的注意力，那就听听他们在谈论什么。我的一个朋友对自己的智商非常不自信。毫无疑问，他很聪明，但是

几乎在任何谈话中，他都要抛出一些信息来证明自己有多聪明。我最后几次见他，他提到他的航空杂志智力测验的表现，一个同事对他说的聪明话随口进行的评价，以及聪明人是如何看纪录片的（噢，他也看纪录片）。这是这个人生活中的主导话题，在与他的所有交流中都很明显。

第二种捕捉人们关注主题的方法是观察他们如何引导对话。几年前，我和两个老朋友一起吃晚饭，一个是富有洞察力的临床心理学家，另一个是建筑师。聊着聊着，这位心理学家不经意地对建筑师说道："听起来你好像在经济上遇到了一些问题。"建筑师对心理学家突然冒出的评论感到非常惊奇。但随后他痛苦地承认，最近由于风险投资他失去了毕生积蓄。心理学家的评论让我感到惊讶，因为我丝毫没有从对话中察觉出我的另一位朋友有经济上的问题。

之后，我问我的心理学朋友，他是怎么知道我们的朋友有经济困难的。他大笑，因为在他看来这很明显。饭桌上有几次，建筑师转移了话题——都和钱有关，类似金融损失或者投资。建筑师和我都没有注意到他换话题的规律。从那以后我就学会了，转移话题标志着说话者心中所想的事情。

◎分析语言以更了解我们自己

如果你可以获得自己一天中使用的所有词汇，你想要了解自己的哪部分？通过语言研究，我可以为自己回答这个问题。至少对我来说，这是非常有帮助的。

我们实验团队研究自然语言的方法之一就是录制孩子、大学生、已婚夫妇和年长者的日常对话。我之前的一个学生马蒂亚斯·梅尔在开发记录设备方面有帮助作用，设备被称为电子激活录音器，或简称EAR。EAR是一个电子录音器，几天内，每12～14分钟可以录音大约30秒。马蒂亚斯现在是一位享誉国际的研究员，他花费了数千个小时在调制EAR上，以便它能够承受佩戴者的折磨。测试阶段的一部分是要求研究的参与成员佩戴几天EAR，然后将录音内容转录为文字。我自己已经做了几次。佩戴EAR的第一周，我的儿子大概12岁。在我的印象中，那周没有发生什么事件——琐事、一次家庭出游，和往常一样。几天以后，我在转录音频时发现，跟儿子说话的方式让我感到难过。我的语气总是很冷、很有距离感——配合他的冷漠。（又或者是他在配合我的冷漠？）在我们的交谈中，我使用大词、高比例的冠词，较少使用代词，尤其是"我"。我跟妻子和女儿说话的时候，语气就更温柔和亲切。

听自己说话和在纸上看到自己的文字这一经历对我产生了深刻的影响。那之后，我有意识地对儿子更温柔些，并在心理上与他更亲近。需要

注意的是，我并不是有意识地改变自己使用代词和冠词的方式。相反，我改变了自己的行为和态度，然后假设功能词会因此而改变。

多年来，我还分析了自己的邮件、课堂讲座、专业文章和推荐信中的语言。有些时候，我的语言表现方式是可预见的，有时就不是；每当不是的时候，我都会更进一步了解自己。确实，每当遇到自己的行为和认知中的自己不同时，我都感到很震惊。这也引出了“我该怎么做”这一问题：改变我对自己的看法，改变我的行为，还是改变我的语言。

这就回到了本书反复出现的主题：语言以何种程度影响着人们的心理状态，或者语言只是反映了人们的心理状态？通过改变使用语言的方式改变我们的心理状态是否可能？例如，如果约翰·克里的顾问建议他使用更多的“我”和更少的“我们”，他是不是就能在2004年的大选中胜出呢？语言上的改变是否会改变他非正式的说话风格和僵硬的肢体语言呢？我对此严重怀疑。人们可以通过训练来改变自己的语言，但没有令人信服的证据表明语言会影响他们的性格、行为或情绪。相反，如果约翰·克里尝试放松下来，变得更真实和诚恳，那他的语言也会随之改变。语言是对一个人的有力反映，但不能改变人本身。

第5章　情绪测定

人们抑郁情绪越严重，越可能在写作或说话时使用“我”字。自杀性诗人在诗中使用“我”的频率远高于非自杀性诗人。

谁说足球比赛没有哭泣？纽约喷气机队前一天以22：24惨败给美洲虎队，教练雷克斯·瑞安早晨起来后当着球员的面发表了慷慨激昂的讲话，情绪激动到自己都流了眼泪……右内边锋达米安·伍迪告诉《邮报》：“他没有痛斥我们，他只是情绪比较激动……他哭了。”“雷克斯非常信任我们球队，我甚至无法用语言形容，如果我们没有抓住机会就会觉得羞愧。”

“看到他那样我有点不安，”后卫达瑞尔·雷维斯对《邮报》讲道，“我也心烦，和他心烦的原因相同。”

当被问到是否曾经参与过情绪如此高昂的会议时，雷维斯回答：“不，我从未参加过教练哭成这样的会议……将来，我希望有更多喜悦的泪水，而不是今早这样。”

——麦克·坎尼扎罗，《纽约邮报》，2009年11月17日

所有迹象都指向本周纽约喷气机队遇到的麻烦，教练雷克斯·瑞安在向球员发表言论时哭了，沉浸在情绪中难以自拔。不知是不是表演，瑞安在用眼泪鼓励队员、团结队员，这正式说明喷气机队大势已去。

——山姆·希契科克，新泽西编辑部（NewJersey Newsroom.com）2009年11月20日

人们情绪很激动时，会吸引我们的注意。一个成年足球教练哭了，对他的球队、对手和看球的人来说是重要的信息。最初《纽约邮报》认为，经历了主教练瑞安的情绪波动后，他的球队准备好去波士顿参加接下来和新英格兰爱国者队的对战。这篇文章发表几天后，知名的新泽西编辑部视这次情绪波动为灾难的前兆。事实是，的确是灾难。周一的比赛中，纽约喷气机队以14：31被爱国者队碾压。接下来的一周，没有与眼泪的相关报道。

情绪改变人们看待世界和思考世界的方式。它可以激励人们努力工作，也可以让他们因失望而放弃。情绪可以拓宽我们的视野或限制我们的视野，使我们一遍遍地思考同一个主题。情绪引导我们的思想，影响我们交谈的方式和与他人相处的方式。我们不仅需要了解自己的情绪，也需要阅读他人的情绪，以便了解他们在想什么，准备做什么。

如果他人在哭泣、喊叫或大笑，这时候阅读他们的情绪通常很容易。其他时候，情绪通过面部表情、语气或非言语行为微妙地传达。然而，很多时候，人们可能感受到了某种情绪但是没有表达出来。我们所有人都有过这样的经历：不知道我们的约会对象、父母、老师、老板或客户是否喜欢我们。在亲密关系中，有些人可能忽略了一些重要的情感线索，这些线索可能破坏了与他人的关系。阅读别人邮件、短信、推文或者信件的过程中，我们大多数人可能忽略了对方想要传达的情感线索。

本章的核心问题就是我们如何通过他人的语言检测出人们的情绪。表面上看，这听起来像一个简单的任务。如果人们是幸福的，他们应该使用开心的文字；如果悲伤，他们应该使用悲伤的文字。如果只有这么简单就好了。通过计数情感词汇来衡量情感是一个不错的开始，但是这些方法忽略了一个核心问题：情绪会影响人们思考的方式。如果我们能够想出一个方法直接衡量人们的思考方式，就能够想出一个研究和理解情绪更好的方法。正好有这样一种方法。功能词在追踪人们思考风格上做得很好。不足为奇的是，这些相同的词汇能够为洞察人们的情感状态提供线索。

不同的情绪，不同的思考方式

一个好的开端是思考三种截然不同的情绪状态：幸福、悲伤和愤怒。我们都同意，这些都是常见的情绪状态，它们拥有自己独特的生理和心理感受。这些不同的情绪使我们看待世界的方式有所不同。当考虑情绪和语言时，去看一下诗人如何写作会很有帮助。毕竟，诗人一生中的大部分时间都在记述他们的情绪反映。诗人在描写幸福、悲伤和愤怒时使用的功能词是否不同呢？来看一些埃德娜·文森特·默蕾的作品，她是第一个获得普利策诗歌奖的女性。默蕾写作于20世纪前半个世纪，她被认为是一个具有自由精神的作家，关于爱情、失败和人际关系的写作非常有影响力。从下面例子中对比她使用词汇的不同。

幸福的例子	悲伤的例子	愤怒的例子
《往事》	《中期》	《巫师－妻子》
我们非常累了，我们那时很愉快，	哎，我很累，我感到困乏	她脸色不红润也不苍白，
我们已经在渡船上来回踱步了一晚上。	太多了——我是个人而已，	而她永远不会是我的；
我们冲着披肩覆盖着的脑袋欢呼：“早安，妈妈！”	我必须睡觉。尽管你又死了，	她从一个童话故事中了解自己的双手，
买了一份早报，我们没人读；	我是个人，我必须睡觉。	情人节卡片上她的嘴唇。
她为苹果和梨哭泣道：“上帝保佑你们！”		她的发量多于所需；
除了我们的地铁费，我们把所有的钱给了它。		无忧无虑曾经对于我来说是灾难！
		她的声音是一串彩色的珠子，
		通向海的步伐。

米莱的例子展示了研究员的发现。当人们记述积极经历时倾向于高频率地使用“我们”。开心的人描述时会更具体，依靠具象名词和特定的时间地点指代词。其他研究表明，积极情绪会改变人们的看法，他们会以更开放的方式看待世界——有时被认为是开阔的和发展的思考方式。悲伤的情绪通常导致人们向内关注。代词倾向于追踪人们的关注点，当人们情感或

身体上承受着巨大痛苦时，他们倾向于使用更多的“我”。和其他多数情绪不同，悲伤与回顾过去和展望未来相关联。换句话说，与其他强烈的情绪相比，当人们悲伤和沮丧时，更倾向于使用过去和将来时态的动词。

虽然愤怒被列为消极情绪，但是愤怒和悲伤完全不同。当人们感到愤怒时，他们关注他人而不是自己。愤怒的人除了使用更多的第二人称代词（例如“你”）和第三人称代词（“他”“她”“他们”）外，还会使用现在时态。

当有事情让我们感到难过或生气时，我们总是想要知道为什么会发生这样的事情。我们会使用认知词汇，认知词汇反映因果思考和自我反省。但对于积极情绪就不是这样，如骄傲和爱。当高兴和满意时，我们大多数人乐于让喜悦冲刷我们，而不会反省。换句话说，消极的感受使我们深思；积极情绪使我们高兴得有点傻傻的。

多数检测短暂情绪的实验研究依赖于大学生的写作，内容与引起幸福、悲伤和愤怒的大事件相关。这些实验研究被视为通往人们感受和表达的线路图。

代词和痛苦：自杀性诗人的语言

一年当中，5%以上的成年人会经历一次抑郁症。抑郁症状与身体健康、社交和工作问题以及更高自杀风险相关联。已知有许多可能影响抑郁症的因素，包括重大的生活动荡、遗传倾向和社交孤立。抑郁症的一个特别突出的理论认为，当人们变得沮丧时，他们往往专注于自己，甚至达到了病态的程度。他们翻来覆去地想自己的焦虑感、悲伤和无价值，同时很少关注周围的世界。

还记得代词反映人们的关注点吧。鉴于抑郁导致人们向内关注，那么抑郁阶段人们会使用更多的自我指示代词，特别是第一人称代词单数，例如“我” 和“我的” 。几项研究发现了这个规律。人们抑郁情绪越严重，越可能在写作或说话时使用“我”字。最令人吃惊的是，对于大学生来说，“我”字是抑郁的更好预测者，而不是消极情绪词汇。

作家抑郁的比例非常高，尤其是成功的诗人。近期研究表明，出版诗人去世时比作家和艺术家年轻，而且自杀率有20%。虽然写诗是一件很有压力的事情，但是更有吸引力的一个解释是，有抑郁倾向的人被写诗吸引，某种程度上，他们在试图了解自己的情绪波动。这对两极型忧郁症，也称躁郁症，尤其正确。两极型忧郁症毒性很强，因为它有明显的遗传基础，通常使人突然处于极端的情绪波动中，且没有明确的起因。与其他类型的抑郁症不同，被诊断为两极型抑郁症的人更有可能自杀。

凯·雷德菲尔德·杰米森是约翰·霍普金斯医学院一位受人尊敬的科学家，她撰写了大量关于艺术气质和两极型忧郁症之间密切关系的文章。通过诗人的回忆录，家庭成员、朋友或传记作者的叙述，以及对诗人诗歌的研究，她发现，不成比例的诗人具有的症状与两极型抑郁症的症状一致。有没有可能通过计算机分析诗人发表的作品察觉两极型抑郁症和自杀倾向？香农·斯特曼现在是一位临床心理学家，我们和他一起检验了18位诗人发表的诗，其中9位选择了自杀结束自己的生命。我们发现，自杀性诗人在诗中使用“我”的频率远高于非自杀性诗人。特别奇怪的是，这两组诗人在使用消极情绪词汇上没有差异。虽然诗人的样本很小，结果却令人印象深刻。

随着更加深入地研究自杀性和非自杀性诗人的语言，我们被一些事实所吸引。在使用“我”字相关代词上，自杀性诗人似乎在心理上与他们的悲伤和痛苦紧密相连，非自杀性诗人则不是这样。

举例来说，看一下西尔维娅·普拉斯那首有名的诗《疯女孩的情歌》

（Mad Girl’Lovesong）中的第一句，她在悼念逝去的爱情：“我合上眼眸，世界倒地死去；我抬起眼帘，一切重获新生。（我想你只是我脑海里的幻象。）”再看一下著名诗人丹妮丝·莱维托芙《婚姻的疼痛》（The Ache of Marriage）中的第一句，对比一下她的悲伤：“婚姻的疼痛：大腿和舌头，亲爱的，过于沉重，它在牙齿间跳动着作痛。”尽管两个诗人谈论的是类似的话题，但是普拉斯使用“我”的方式表明她在拥抱她的损失。而另一方面，莱维托芙似乎将痛苦拦在双臂以外——仿佛她在以一个更远的第三方视角来看它。确实，如果一个人阅读这两位令人印象深刻的诗人的作品集，会发现两人拥有损失或者说拥抱损失、疏远和沮丧的方式明显不同。普拉斯可能因为这个原因才更受欢迎。通过第一人称代词单数，她将我们拉至边缘，使我们能够感受到她的绝望。

傲慢、失落和沮丧：朱利安尼市长和李尔王

与悲伤和沮丧紧密相连的感受是失落感，通常来自失败和被拒绝。2000年，《纽约时报》的一篇头版文章报道称，新闻界成员目睹了纽约市长鲁道夫·朱利安尼的性格发生一些明显的变化。我个人的经验是，人的性格通常不会改变，《纽约时报》的这则报道引起了我的兴趣，让我想要进行更深入地研究。

在鲁道夫·朱利安尼做市长的八年里，媒体界对他的描述各式各样，有时认为他是反应迟钝的欺凌者，一个充满愤怒和自以为是的人，有时认为他是一个充满温暖、富有魅力和热情的人。虽然朱利安尼在改变，但这些相互矛盾的评论通常出自同一批人。大多数纽约人比较赞同的一点是，他是一个有效率的市长。他帮助经济救市，减少了犯罪，并恢复了旅游业。因为市长当得很成功，他参与了2000年的美国参议院席位竞选，和希拉里·克林顿是对手。

2000年年末，朱利安尼的生活在两周之内发生了翻天覆地的变化：他被诊断出前列腺癌，退出了和希拉里·克林顿的竞选，在国家电视台上宣布和自己的妻子分开（妻子不知情的情况下），几天后，承认他和朱迪丝·内森的"特殊朋友关系"，并在之后娶了这个人。到5月中旬，他一直在接受癌症治疗，住在一个朋友的公寓里。到六月上旬，朋友、熟人、老对手和媒体都注意到朱利安尼好像比以前更真诚、更谦逊、更热情。

一个可靠的抑郁症预测点就是经历痛苦的生活事件。确实，一个人在一段时间内经历的创伤性起伏越多，生病和患抑郁症的可能性就越大。我们能够从朱利安尼的语言中看到他的性格变化吗？幸运的是，纽约市长参与过很多新闻发布会，我们可以拿来分析。具体来说，我们想知道，与早期在位时相比，他经历情感波动时期使用功能词的方式是否有变化。

与任职市长第一年相比，朱利安尼使用"我"的频率急剧上升，大词频率有所下降，积极情绪和消极情绪词汇都有所上升。他还少用了第一人称代词复数"我们"。我们在前面的章节中提到，自大、情感疏离和地位高的人更常使用"我们"。男性尤其会使用"我们"的疏远含意和皇室用法："我们需要分析那些数据"或"我们不会忍受更高的税收"。朱利安尼的情况是，他的语言显示他的性格发生了有趣的变化，从冰冷、有距离的人转变成一个更热情和更直接的人。

当朱利安尼项目的第一阶段完成时，出现的结果异常熟悉。然后，我突然想到莎士比亚的《李尔王》。在剧中，李尔王开始是一个傲慢的统治

者，要求他的女儿们公开表达对他的爱和钦佩。他最喜欢的女儿考狄利娅拒绝这样做，并且最后离开英格兰，嫁给了法国国王。战争、斗争、指责，然后痛苦随之而来。（注：这是CliffsNotes的版本。）在最后一幕中，身受重伤的国王李尔面对心爱女儿的尸体，他变了。面对失去的创伤，他的个性散发出温暖和人性。看出和朱利安尼的联系了吗？请阅读李尔王第一次和最后一次的台词。

第一幕，第一场　李尔王宫中大厅

李尔：告诉你们吧，我已经把我的国土划成三部；我因为自己年纪老了，决心摆脱一切世务的牵萦，把责任交卸给年轻力壮之人，让自己松一松肩，好安安心心地等死。康华尔贤婿，还有同样是我心爱的奥本尼贤婿，为了预防他日的争执，我想还是趁现在把我的几个女儿的嫁奁当众分配清楚……告诉我，你们中间哪一个人最爱我？我要看看谁最有孝心，最有贤德，我就给她最大的恩惠。

第五幕，第三场　多佛附近英军营地

李尔：你们都是些石头一样的人；要是我有了你们的那些舌头和眼睛，我要用我的眼泪和哭声震撼穹苍。她是一去不回的了。一个人死了还是活着，我是知道的；她已经像泥土一样死去……一场瘟疫降落在你们身上，全是些凶手，奸贼！我本来可以把她救活的；现在她再也回不来了！考狄利娅！等一等。嘿！你说什么？她的声音总是那么柔软温和……从前我一举起我的宝刀，就可以人把他们吓得抱头鼠窜；现在年纪老啦，受到这许多磨难，一天比一天不中用啦。你是谁？等会儿我就可以说出来了；我的眼睛可不大好……请你替我解开这个纽扣；谢谢你，先生。你看见了吗？瞧着她！[1]

这两段对话的分析和朱利安尼市长身上的改变形成了惊人的一致。事实上，朱利安尼和李尔早期傲慢阶段代词和大词的使用与创伤后热情诚恳阶段相对不同，使用方式令人困扰。在傲慢阶段，李尔和朱利安尼使用“我”字和情感词汇的频率较低，与此同时，使用“我们”和大词的频率较高。这些使用模式在两个人遇到人生重大改变（李尔的情况是生命结束）时发生了逆转。生活模仿艺术，而科学记录这一切。

2000年朱利安尼遭遇个人危机后，他的故事以另一种有趣而重要的方式打开。在他的个人危机发生一年多后，“9·11”事件造成了位于世界贸易中心的双子塔倒塌，死亡人数接近3000人，此时是朱利安尼任职期的最后一个月。据大家所说，朱利安尼成为美国和纽约市一位强大和富有同情心的领导者。

恐怖袭击事件后第一周的新闻发布会上，朱利安尼表现出真挚的热情和气度。对他的语言分析展现了一种新的词汇使用模式。他使用“我”和“我们”的频率稍高。但是，他使用“我们”的方式非常突出。在他早期任职期间，他的“我们”通常指代含糊，指代整个社会。袭击事件后，朱利安尼的“我们”更与个人相关并更有针对性，指代纽约市民或政府中的特定组织。

朱利安尼项目补充了自杀性诗人的结果，显示出情感状态和功能词间的联系，尤其是代词。情感状态可以反映和影响我们与他人的社交关系。就性质而言，代词反映了说话者和沟通者之间的关系。代词和其他隐性功能词可以作为微妙情感探测器，我们大部分人无法有意识地识别。

[1] 两段台词皆为朱生豪译本。——译者注

创伤如何显露：语言是窗口

人在处理情感痛苦方面至少有两种方式：承认它和躲避它。自杀性诗人、小说中的国王和真实的市长都承认他们的痛苦和失去。社交方面，“我”字使用频率升高，使他们显得更加内省和脆弱。向内寻找会加重痛苦，激励人们理解和忍受痛苦，并提醒他人注意他们的情感痛苦。

另一种常见的处理痛苦的方式是躲避它，或者从某种意义上说，使自己远离它。回想丹尼丝·莱维托芙的《婚姻的疼痛》，她在总结过往经历时，使用了一种与个人无关的远离方式。其他躲避策略还有试图将不想要的情感经历从脑海中全部清除。事实上，远离痛苦的方式是调节情感的有效方式，尤其是在短期内。假设我在一次会议之前得知我的狗去世的坏消息，那我理应先把自己的情绪放在一边，继续把会开完，仿佛什么都没有发生一样，即使我内心是崩溃的，很想趴在地板上痛哭一场。

短期内躲避情感痛苦似乎是一种艺术形式。在一系列成功的实验研究中，丹·韦格纳和他的同事们证明，人们无法停止对情感事件的思考。确切地说，人们需要关注一些其他东西。如果在会议之前发生了可怕的事情，韦格纳会建议你关注会议，不要对自己说：“不要想狗的事情，不要想狗的事情。”

人们什么时候可以在面临创伤经历时自然地使用躲避策略而不是承认策略？直到最近，科学家们才能够在人们表达创伤时追踪他们的反应。很明显，大多数人处理剧变时倾向短期内同时采用躲避策略和承认策略。

◎创伤震惊后的第一个小时

几年前的假期，我和家人住在朋友赫克托家里。他让我听他的语音信息，如果有事就联系他。我们待了几天后，收到一条男性的信息，语气轻柔无起伏。

赫克托，我是诺兰。发信息是想告诉你，玛格丽特昨晚去世了。几天前她的情况急转直下。谢谢你上周打给我。非常感谢。周一会有一个追悼会。晚些时候会给你打电话。保持联系。保重。再见。

我不知道诺兰和玛格丽特是谁，也不知道他们的关系。但很明显，诺兰已经完全崩溃了。当我再次听他的语音时，我在想，我是如何知道他很崩溃的呢。他也没说自己很难过或很痛苦。他没有哭，声音也没有颤抖。他的语言没有任何修饰，但是，我却不太习惯。

作为代词的鉴赏家，我很震惊，诺兰在信息中没有使用任何“我” 或者“我们” 。事实上，在听过诺兰语音留言后的几年里，我听过至少三个人打电话来告知亲近的人的死讯。和诺兰一样，他们也很少使用“我”字。

最近，我把博客上关于发帖当天家人、伴侣或兄弟姐妹去世的博文的第一句话编排成册。与他们该博文的上一条的语言进行对比，同样的模式出现了。在最痛苦的那几个小时，大多数人使用“我”的频率相对很低，使用消极情绪词汇的频率较低。他们的语言相对简单，使用词和句更短，使用认知词汇更少。

在经受创伤性损失的当下，人们通常感到迷茫、麻木并极其痛苦。作为减轻痛苦的一种方式，他们会调节自己的注意力，从自身转移开。强烈的悲痛使人们较少关注自己和自己的情绪。相反，他们更多关注逝去的人、家人和死亡的细节。

追踪一个国家的情绪：2001年9月11日的博客

在听到情感上压倒性的消息后，人们从心理上让自己远离创伤是很合理的事情。当大批人群同样面临创伤经历时，这样的事情会在更大范围内发生吗？在美国，当人们得知亚伯拉罕·林肯和约翰·肯尼迪遇刺的消息时，得知1941年的珍珠港事件或者“9·11”事件时，他们会回想自己在哪里、在做什么。当人们得知一件民族性的共有创伤事件时，情感上也会像遇到个人事件一样远离自己吗？

与早期民族动荡不同的是，“9·11”事件发生时，正是大批人开始使用博客的阶段。迈克尔·科恩是我的研究团队里的一个本科学生，他掌握了大量的计算机专业知识，在袭击事件发生几周后，他来到我的办公室，提议分析上千个博客文本，以观察人们在事件发生前和事件发生后的几周和几个月内的写作和想法是如何变化的。

那时我们与一个流行的博客网站（LiveJournal.com）有合作，我、迈克尔和马蒂亚斯·梅尔保存了上千人的帖子，他们每周发帖三到四次，我们保存了他们在“9·11”事件前后各两个月的帖子。我们选择发博文的人居住在美国，年龄范围广泛。据各方面信息所知，博文主人是普通人，他们单纯地喜欢在博客上发表各类主题的文章。7万多条的博文分析显示，“9·11”事件前、事件中和事件后，代词和情感词汇的使用发生了惊人的改变。

和得知死亡信息不久后发的电话通知一样，线上博客使用者在他们得知“9·11”恐怖袭击后，使用“我”的频率立即下降。从图5-1中可以看出，“我”的使用频率从6.2%的基线急剧下降。从数据角度讲，这是一个令人瞠目结舌的惊人改变。

抓紧你的座椅，后面还有。“我”的频率下降的同时，第一人称代词复数“我们”的使用频率跳至一个更高点。正如图5-2显示，事件发生后“我们”的使用频率几乎是事件发生前的两倍。回忆一下前面章节提到的“我们”的几种不同用法。人们使用的“我们”是混合的“我们”，既指美国人民也指家人。

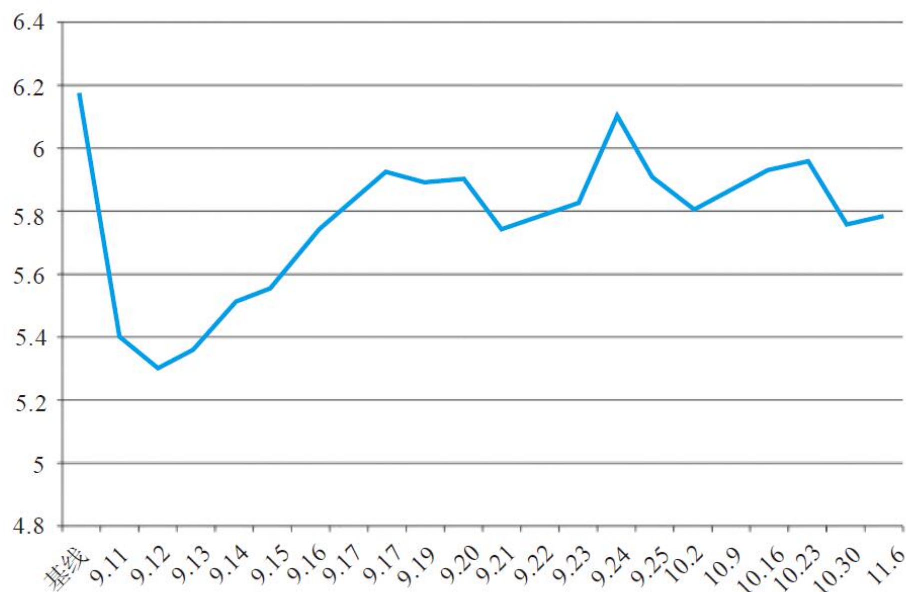

6.4
6.2
6
5.8
5.6
5.4
5.2
5
4.8
基线
9.11
9.12
9.13
9.14
9.15
9.16
9.17
9.17
9.19
9.20
9.21
9.22
9.23
9.24
9.25
10.2
10.9
10.16
10.23
10.30
11.6

图5-1 “我”

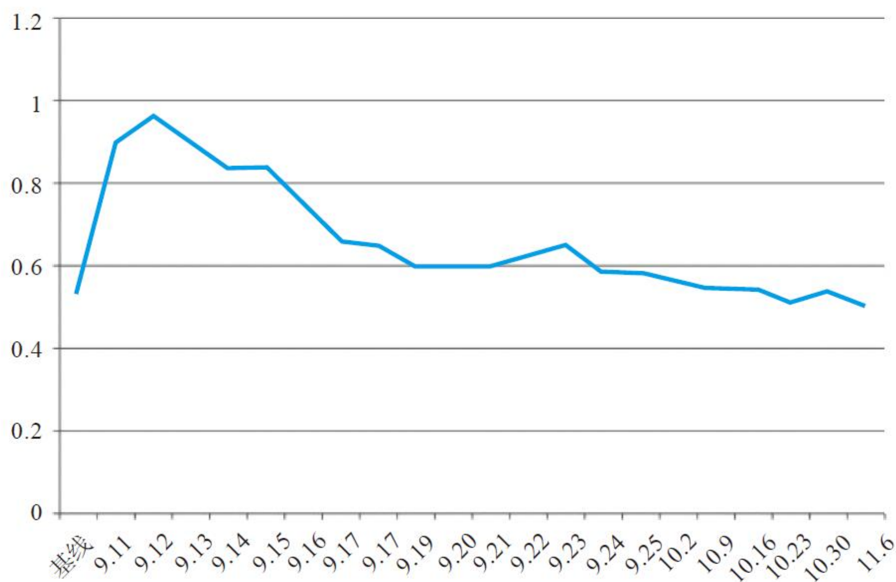

1.2
1
0.8
0.6
0.4
0.2
0
基线
9.11
9.12
9.13
9.14
9.15
9.16
9.17
9.17
9.19
9.20
9.21
9.22
9.23
9.24
9.25
10.2
10.9
10.16
10.23
10.30
11.6

图5-2 “我们”

注：图中表示“我”（上）和“我们”（下）的百分比，数据基于1084位博客使用者在2011年9月11日附近两个月内发的博文。

下面是两个例子。

一个25岁的女性，她的上一条博文讲的是她对一位男士有好感，以及撞见自己前男友很尴尬，现在她被“9·11”事件深深震撼。

我眼看着大楼倒塌，世界贸易中心倒下的时候我哭了出来……愤怒取代了我的……恐惧。一想到我们的家不再安全。我想失去安全和失去生命都让我们感到愤怒。

这个男性之前的博文都与哲学和文化有关。事件发生前一天，他最后写的文章有关艾茵·兰德和自由主义。第二天他成为社会责任感的化身。

我们可以做些什么？血库现在可能爆满，也许未来几天都会爆满。不要因此你就停止。国家需要你。你可以有所为。

大约92%的博文在事件发生24小时内提到了恐怖袭击。一天之内，它迫使人们拥抱亲人、邻居和同胞。

博客中情感词汇的使用支持了代词的分析结果。在积极情绪词汇相对短暂下降和消极情绪词汇相对短暂上升后，人们使用的情感词汇回归常态。而且，事件发生前他们更倾向于表达积极情绪。为了捕捉到这些结果，我们的电脑程序计算出反映积极情绪的词汇（例如“爱”“开心”“礼物”）和反映消极情绪的词汇（例如“仇恨”“哭泣”“担忧”）。正如图5-3所示，在事件发生前两个月内，人们使用积极情绪词汇的频率远高于消极情绪词汇。基本上人们与他人的沟通是相当积极的。

7万多条的博文分析显示，恐怖袭击导致消极情绪表达在两天内急剧上升；然后消极情绪表达在11天后回到事件发生前的水平。与此同时，积极情绪词汇的下降呈现出更惊人的模式。积极情绪词汇在“9·11”事件后发生急剧下降，然后在四天内回到了事件前水平。事件发生10天后，人们的积极情绪词汇使用量高于事件前水平。

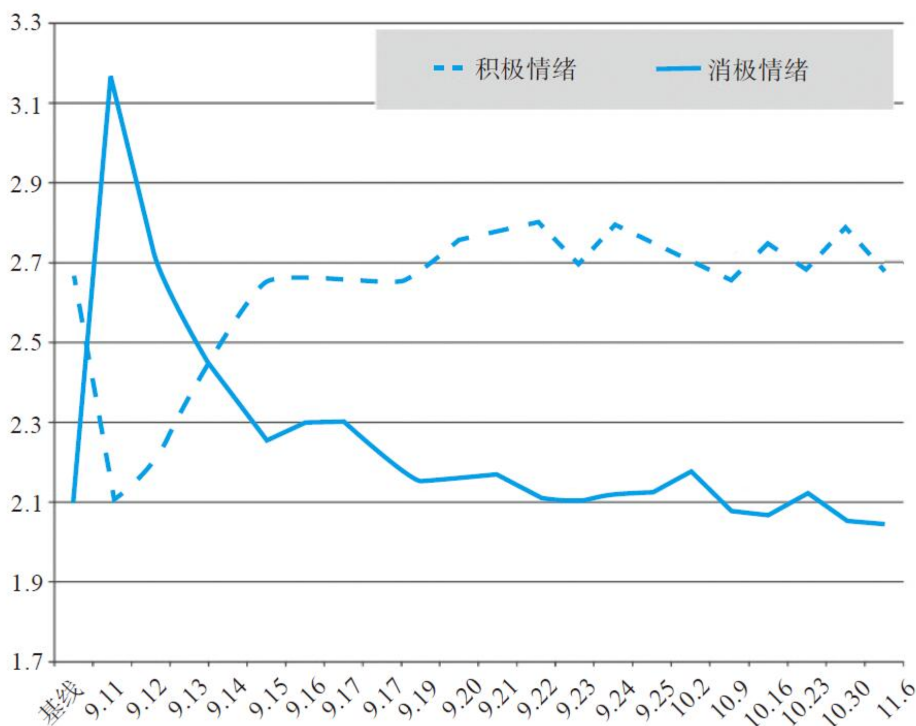


图　5-3

注：曲线表明了1084位博客使用者使用积极情绪和消极情绪词汇的比例。基线是“9·11”事件前两个月的博文的平均值。9月25日之前的数据是按天计算，之后一直到11月6日是按周计算。

“9·11”项目的最后一项重要发现。事件发生后的五六天，博客使用者使用认知词汇的频率远高于事件发生前。回忆一下认知词汇包含反映因果思考（例如“因为”“引起”“影响”）和自我反省（例如“理解”“意识”“意义”）的词汇。认知词汇通常表明，人们试图理解他们生活中正在发生的事情。

意想不到的事件发生后，认知词汇使用量立即上升是合乎情理的事情，因为我们都想知道发生了什么以及为什么会发生。然而，在“9·11”事件发生大概一周后，博客上使用认知词汇的数量开始下降，在接下来的两个月，认知词汇使用量降到前所未有的至低点。思考标记水平远低于“9·11”袭击事件前。

考虑一下所有这些结果意味着什么。对上千博客使用者的分析描画

出，在经历大的情感波动后，通常情况下人们每天的想法是怎样的。值得注意的是，在“9·11”事件发生几天后，大多数博客使用者回归记述他们的日常话题——购物计划、男朋友、色情文学、宠物，和往常一样。然后，总览所有这些话题，相同的模式出现了。对博客使用者的语言总结如下。

·共有创伤把大家团结在一起 。他们更多地关注他人，并视自己为大家庭的一部分，这可以从温暖的“我们”，即“你和我”形式的“我们”使用量上涨看出。

·共有创伤使人们将注意力从自身转移 。即使人们可能感到难过，但是他们不会沮丧。回想一下，真正难过的人使用“我”的频率较高。

·共有创伤在许多方面是积极的经历 。“9·11”事件后至少两个月内，人们更多地表达积极情绪，社会连接比事件发生前更紧密。

·共有创伤使人们更笨 。好吧，也许不是更笨，但是肯定更不善于分析。事件发生一周内，人们写作方式更简单，表明他们没有很深入地思考自己写作的主题。他们看上去更消极，仿佛正在接受新信息。

·人们对于创伤经历的反应随时间而改变 。人们思考、感受和关注自己世界的方式在经历情感波动后的几小时、几天和几周后会发生很大改变。

“9·11”事件之外：天灾和人祸

进化心理学家看到关于“9·11”事件的发现，指出它们具有进化意义。如果我们的一个小团体在热带草原上遭遇到另外一个团体的袭击，那么我们团结起来，一致对外，为未来可能的袭击做好准备是非常重要的。这些属于适应反应，可以帮助提高个人和团体存活的可能。

共同经历情感波动后，是否存在不会直接威胁团体的相同语言模式？答案显然是肯定的。在我的职业生涯中，我已经研究了一些大规模动荡的社会影响和心理影响，包括1980年华盛顿州圣海伦火山喷发、1989年旧金山海湾地区洛马普列塔地震、1991年第一次海湾战争、1997年戴安娜王妃去世，以及1999年12名学生在得克萨斯州农工大学惨死，期间正在建造他们的传统篝火。

所有这些研究中，相同主题的出现支持了“9·11”的发现。最令人吃惊的现象是，所有类型的动荡都会将人们团结在一起。人们变得更无私、更关心他人、更主动地寻求与他人的联系。圣海伦火山在他们社区倾倒了几乎10厘米的沙状火山灰，几周后我采访了华盛顿州雅吉瓦人，大多数居民称，那是一次可怕的经历，但尽管有损失，他们还是很高兴它发生在他们的生命中。大多数人说，他们和从未了解过的邻居见面和聊天。

得克萨斯州农工大学篝火倒塌，12名学生死亡的事情轰动了大学社区。此外，在采访、网上博客和报纸报道中，社会语言是无私和热情的。事实上，大学的社会连接最终变得比以往更加紧密。第二年，在跟进学生的身体健康时我们发现，这一年因为生病去健康中心的学生数量比篝火灾难发生前的那一年少了40%。农大的邻居，得克萨斯大学奥斯汀分校没有发生相同的情况。

可怕的经历能显示出最好的我们，虽然承认这件事几乎是异教的行为。创伤的本质可以摧毁生命，也可以丰富人生。

使用语言分析来指导情绪波动后的心理健康治疗

心理健康界隐藏最好的一个秘密是，实际上当大多数人面对真正可怕的生活事件时会应付得很好。这与我们很多人认为的正好相反。当研究人员问我们，如果可怕的事情发生在我们身上，我们会如何应对时，我们倾向于认为，面对个人创伤，我们在心理上会无法承受。事实上，大多数经历过折磨、强奸的人，或从车祸、空难、其他难以想象的事件中幸免的人不会出现创伤后应激障碍（PTSD）症状或重度抑郁症状。

看来，人类的适应能力是极强的。他们同样天生知道当创伤发生时要做什么。要知道，绝大多数被研究的博客使用者心理都很健康，这意味着，不管他们在做什么，都应该是个不错的选择。有个发现是明确的，有点违反直觉：使自己的注意力从痛苦中转移是健康的。花钱、帮助他人、玩游戏或者收拾房间都是由来已久的应对极度令人不安消息的方法。如果你想和别人聊这件事，那就跟他们聊。如果你想做自己，那就做自己。没有可靠证据证明哪个单一的应对方法可以适用于所有人。

如果你是一名心理健康专家，这个建议是无法令人满足的。如果一场灾难发生，你在电视上看到人们深处苦痛之中，你的自然倾向是跑过去帮助他们。此外，所有证据表明，你能提供帮助的最好方式是确定人们的基本需求已经得到满足，确定他们可以回答你的问题，他们会听你的提问。人们在经历灾难后的短时间内不需要深度心理治疗。非常不需要。

20世纪90年代，一种善意的心理干预流行起来，它鼓励心理受到创伤的人们在创伤发生后的72小时内说出内心最深处的感受。这种活动后来被称为集体晤谈（CISD），听起来非常合理，并且被应急人员、大型企业和全世界的政府派别采用。虽然这个项目的初衷是好的，但是它有很多严重的问题。经过一系列研究，现在基本一致认为，集体晤谈可能弊大于利。通过观察创伤后一小时人们在网络上使用语言的方式可知，在情感动荡事件发生后不久进行内心情感调节不是健康人做的事情。

大多数人在经历波动后需要社会支持。我和我的学生在研究博客时发现的最惊人的一个现象是，人们寻求和接受他人支持的程度。人们在经历波动感到沮丧时，只会提到他们的情感状态，随后会收到来自朋友和陌生人的一堆美好祝愿。

辛迪·钟的博士论文涉及饮食博客的语言，连续几个月，她通过饮食博客社区中几百个节食者来追踪什么因素促使减肥成功。很多博客使用者关注减肥细节，但是大多数人也记述他们的人际关系、生活经历和情感问

题。很多减肥专家建议，减肥最好的办法是着魔似的列出每天你所吃的食物、摄入的卡路里和你的运动量。根据钟的研究，事实并非如此。相反，她发现，成功减肥最好的预测器融在社交网络里。也就是说，一个人收到的评论或发出的帖子越多，他（她）减肥成功的概率越大。另外，记述自己个人问题和情感问题的博客使用者减肥成功的概率远高于那些只记述自己饮食的人。

创伤的长期影响：语言如何治愈

根据博客数据显示，大多数人最初应对创伤经历的方式是找到朋友和家人。人们谈论或描述波动的方式有很大不同，可能不涉及任何深层次的情感讨论。

如果人们总是能够按照他们想要的或需要的方式回应情感事件，那么对心理治疗师和医师的需求量可能会减少。创伤，即使是高度个人化的创伤，也是深度社会化的。如果我被强奸或抢劫，我可能不愿意告诉别人，因为我有羞辱感或害怕别人会因此对我不同。所以，我想要公开讨论这些经历的天然反应会被扼杀。同样，我可能需要表现得很坚强、很快乐，以此避免家人的担心和难过。

◎秘密的危害

可耻的、尴尬的或可能损害自己声誉的事件往往会被保密多年。这是在我的职业生涯早期发现的，当时我的一个问卷调查中包含一题是询问人们在17岁之前是否经历过创伤性性经历。在上千份对成年人的调查中，大于22%的女性和11%的男性表示他们曾有过这样的经历。特别引人注目的是，与没有创伤经历的人相比，这些人的健康状况很糟糕。后续研究表明，问题在于性创伤总是秘密的创伤。任何类型的大波动，如果人们隐瞒于他人，都会危害他们的身体健康和心理健康。

隐藏大的情感波动有害是有原因的。我进行首批面谈的其中一个案例说明了这个问题，她是一名35岁的女性，我暂且叫她劳拉，在她12岁时，她的母亲再婚了。婚后的前几个月，劳拉的新继父乔克会时不时地半夜潜进她的卧室抚弄她。她很害怕，试图让他离开，即使乔克一副无所谓的样子。这种情况断断续续地持续到劳拉15岁，她离开家和一个阿姨住在了一起。她描述道：

“我一直和妈妈的关系很好。离婚几乎要了她的命，而她和乔克在一起时那么开心。如果得知乔克对我的所作所为，她一定会很伤心。我是那么想告诉她。你知道生活在一个这样的家庭是什么感觉吗？每天早上醒来，我都会看到乔克和妈妈一起从楼上下来。他微笑着，很友善，好像什么都没有发生过。我恨透了他，却不能告诉别人为什么恨他。每天早上，每次看到那个混蛋，我都恶心地想吐。

回头看这一切，最糟糕的事情是我不再能和妈妈聊天了。我必须在我们之间树立一道墙。如果我不够小心，墙就会坍塌，我就会告诉她一切。对待我的朋友也是一样。我和女性朋友们出去玩时，我们会高兴地聊着男

孩和约会。她们的笑发自内心，而我的不是。如果她们知道我卧室里发生的事情，她们会死掉。”

在过去25年里，我听过或阅读过上百遍这一主题的各种故事。在劳拉的故事里，继父的抚弄是可怕的，但更可怕的是附加影响。她所有的和家人朋友的密切关系都遭到破坏，她的身体健康恶化，而且事情发生的几年内她都无法和任何人谈起这段经历。

不谈论大的情感波动违反了我们的自然行为。正如我们所看到的，情感事件引发对话。如果我们目睹了可怕的事故，发现我们最喜欢的球队赢得了一场重要比赛，得知我们最亲密的朋友刚离开了她的丈夫，我们需要讨论这些事。从根本上讲，想要讨论意想不到的事情和令人不安的事件是人类的天性。在所有被研究的文化中都是如此。人们通过讨论情感经历来更多地了解它们。谈论，是我们理解复杂经历的首要方式。相反，当人们无法谈论情感事件时，他们会想这些事，甚至执着于或反复地思考它们。

将秘密转化为语言：表达性写作

正如第1章所述，秘密和疾病之间的联系促使我开启这个想法。如果我们鼓励人们谈论或记述不安的经历，他们的健康是否会有所改善？结果显示，答案是肯定的。在我们的第一个实验中，大学生参与了写作研究，开始时要求他们连续4天每天按指定话题写作15分钟。参与之后，一半的人被要求4天里记述他们深刻的创伤或压力经历。另一半学生被指示写一些无关痛痒的话题，例如描述物体或事件。

总的来说，那些被要求记述情感话题的人的健康状况比那些记述无关痛痒话题的人要好。6个月后，实验组记述创伤经历的人去医院的比例是对照组的一半。后续实验得到相似的模式。记述创伤经历改善了人们的身体和心理健康。

其他研究员很快发现，表达性写作影响免疫功能和其他与健康和疾病相关的生理过程。来自雪城大学的乔希·史密斯和他的同事们发布了一项有关关节炎和哮喘病人的有力研究，研究显示，写作影响疾病的进程。其他项目中患有艾滋病、癌症、心脏疾病、抑郁症、囊性纤维症，以及一系列其他生理和心理问题的人也从表达性写作中获益。现在，将近25年后，超过200多篇科学报道已经发布，指向写作的力量。

早期的写作研究是开发文本分析方法的动力，文本分析方法可以帮助我们理解为什么写作会起作用。唤起人们对之前波动的感受，迫使他们从不同的角度思考问题。探究情感话题要求人向内看（正如我们看到的，他们使用的第一人称代词单数增多）。它帮助人们整理思绪，并为他们的生活构建更有意义的解释或故事（认知词汇使用量增多）。通过观察人们记述创伤的方式，我们正在目睹他们的思维是如何改变的。

写作不是万能药，而且其影响是有限的。比如，没有任何证据表明情感波动发生后立刻记述会有所帮助。虽然持续几天每天写作相对短的时间总的来说是有效的，但是长期记日记是否一定有效目前尚不清楚。事实上，关于一个特定问题记述的过多也许是一种沉思。我的建议是，如果你对表达性写作感兴趣，可以尝试几天。如果有益处，那很好。如果没有，试试别的。

两者结合：情感和思考方式是硬币的两面

还记得雷克斯·瑞安吗？纽约喷气机队的教练，他在球队输了一场重要比赛后当着球员的面哭了出来。他的情感表现反映出他对球队和球队潜力的看法转变了。同样重要的是，他的眼泪对球员是强有力的社交信号。你应该还记得，其中一个球员说过，教练对球队的付出给他留下了深刻的印象，并且他的哭泣使球队团结在一起。在另外一篇文章中，对于一个体育记者来说，同样的哭泣场景被认为是纽约喷气机队将要分崩离析的信号。

情感不仅仅是对事件的反应。不同的情感会改变我们的思考方式，影响我们回应他人的方式。情感是极度社会化的，因为它可以拉近我们，也可以疏远我们。情感还是有意义的信号，传达人们的动机、目标和意图。功能词和情感状态的密切联系是自然连接的。情感使我们以不同的方式思考世界，而功能词反映了思维的改变。

几个世纪以来，思维和感受之间的关系一直是哲学和心理学热议的话题。亚里士多德和柏拉图都认为，从根本上讲，逻辑和情感是不同的过程。17世纪的作家笛卡儿称，情感破坏了人们理性思考的能力。早期的美国心理学家威廉·詹姆斯也强调，情感和热情常常会蒙蔽人们的判断力。弗洛伊德认为，根本的情感问题是人格和行为的驱动力。

我们开始对情感和理智产生了不同的看法，一部分源于大脑科学的发现。其中对这个新角度最有说服力的一个代言人是神经系统科学家安东尼奥·达马西奥，他研究并记录了遭受脑额叶损伤的人的行为。额叶集成了原始情感中心以及抽象推理和语言相关区域的信息。许多连接非常广泛，以至于明确区分情感和思想是没有意义的行为。

在《笛卡儿的错误》（Descartes'Error）一书中，达马西奥描述了人们玩一种纸牌竞技游戏的过程。健康的、脑部没有受损的人在做决定时对奖励和惩罚高度敏感。而那些额叶受损的人似乎忽略了失败的感受。他总结，失去的情感帮助人们表现得更理智。情感体现思维。

我们的感受影响我们思考世界的方式，这一信息是本章的关键信息。我们的情感影响我们的思考，这反映在我们使用功能词的方式上。引申开来，功能词可以让我们初步了解他人的想法和感受。它们还充当微妙公告，提醒他人我们自己的情绪状态、我们的思维模式，以及我们的关注点。

第6章　说谎的语言

“我”代表无辜。真正有罪的被告使用更多的第三人称。他们试图将罪责从自身转移给他人。

测谎专家已经得知，人在撒谎时会伴随特定的生理反应。真实告白的生理反应是什么？在我的表达性写作研究早期，志愿者来到实验室后会被要求描述一件影响较大的情感事件，同时佩戴测量他们血压、心率、呼吸、肌张力和皮肤电传导（有时称为皮肤电反应）的传感器。他们在描述创伤经历时，通常是独自对着录音带讲。一个学生对我来说印象深刻。下面是他讲述的编辑后的版本。

12月20日晚上10:35，我的父亲向北开车行驶在27号州际公路上。他开着一辆1990年的别克名使，时速大概是65千米每小时，这时一只鹿跳到了他的车前。他停不下来。他的车突然转向，然后失去了控制，导致车身翻转了三次，最后撞到了一棵树上。根据验尸官的报告，他应该是当场死亡。我母亲在凌晨00:15接到电话，然后告诉我这件事。虽然当时只有14岁，但我已经可以应对死亡，并且这件事的影响很小……

这个学生在录音里的语气很平淡，异常冷漠，就像他的语言一样。然而从生物学角度讲，他表现出巨大的冲突和压力。在谈论这件事时，他的皮肤电传导和血压上升，面部肌肉紧张。然而，在项目结束后的调查问卷中，他说道，谈论创伤事件完全没有不安或紧张的感觉。

自我欺骗的语言

我很少见到自我欺骗如此明显的例子。这个学生完全认识到事实，但是他在描述这件事时没能承认这件事给他带来的情感影响，我猜想，在他父亲死后数月或几年后他才能够承认。他的事情使我想起其他涉及表达性写作的研究，实验中临时参与者会记述可怕的创伤经历，但是他们不会提及消极情绪和感受。事实上，那些无力承认他们面对不安经历时的情感反应的人，几乎不会从表达性写作中获益。那些探索自己过往时能够诚实面对的人会发现写作的最大价值。

自我欺骗的代价有点儿争议。大部分现代宗教和心理疗法是基于“你的自我是真实的”这一前提的。有点儿讽刺的是，莎士比亚著作中的名句是欺骗者波洛涅斯说给他的儿子欺骗者雷欧提斯的，欺骗者雷欧提斯最终被欺骗者哈姆雷特以欺骗的方式杀死。不过，自我意识与更好的心理健康和生理健康相关，这一想法有很长的历史。这是有道理的。了解自己的人应该能够更好地衡量自己的长处和做决策时的局限性。

另一种看法是，对自己抱有积极幻想会走向更幸福的生活。如果马娅·安杰卢真正了解成为世界知名作者和诗人的概率是极其渺茫的，她还会去做吗？如果菲尔·赫尔姆斯知道赢得世界扑克大赛的概率小于万分之一的话，他还会参赛吗？（赫尔姆斯赢得比赛11次。）如果杰克叔叔知道彩票中奖的概率是几百万分之一的话，他还会继续买彩票吗？（会的。而且他从来没中过奖。）从一开始，人类不得不尝试不太可能取得成功的事情。人类对自己的能力有着自我欺骗的信念，他们被这种信念所驱动。数据统计，几乎没有人成为世界知名诗人、扑克比赛冠军或彩票中奖者。然而，有一小部分人在这些领域获得了成功，他们的成功通常加剧了我们的幻想。

对我们的能力、人际关系和世界抱有积极幻想可以更安心、压力更小。不好的一面是，我们对自己过度自信是否会严重歪曲事实，产生灾难性后果。例子当然有很多。在罗伯特·特里弗斯一本有趣的书《愚昧者的愚昧：自欺与欺骗背后的逻辑》（Deceit and Self-Deception）[1] 中，他详细描绘了自我欺骗的严重结果，从空难、失策的战争到全球经济萧条。

自我欺骗有很多种方式。那个谈论自己父亲死亡的学生就是一种，人们会否认或无法承认某件事的情感影响。另外一种方式是对自己的能力或处境过度自信。还有一种是坚信一个已经证实错误或无法证实的信念。例如，一个男人在自己的前女友结婚并发出禁制令后依然坚信她还爱他。极端情况下，患有严重精神紊乱症（如精神分裂症）的人的幻想就是这种形

式的自我欺骗。

自我欺骗可以由语言捕捉吗？某种程度上说，可以。回头看一下那个学生父亲去世的故事，三个语言特征跳了出来。

·非个人化的语言。 大多数人在记述个人的情感波动时，对待那段经历会比较个人化。他们会使用像是“我看到”或“我觉得”这样的短语。请注意，这个学生是一次“我”都没有使用过的。

·缺乏情感。 尽管他描述，发生这件事时他才14岁，但是这个学生几乎没有使用情感词汇，尤其是消极情绪词汇。事实上，他少有的情感词汇暗示这段经历是“可以应付的”和“影响很小的”。

·语言具体、生硬和异常冷漠。 你可以发现，他倾向于使用高频率的具体名词（可以通过他使用的冠词来衡量）。他还大量使用动词，尤其是“应该”这样的词。“应该”这样的词在现实事件和个人认知之间产生了一种距离。

◎文学人物中的自我欺骗

自我欺骗的各种形式经常出现在文学作品中。想一下狄更斯《圣诞颂歌》（A Christmas Carol）中的埃比尼泽·斯克鲁奇。故事的开始，斯克鲁奇是一个冷漠的老头，他对圣诞节、家庭和亲密的人际关系持轻蔑态度。斯克鲁奇的出场是在平安夜前夕，他在办公室内苛责他的店员鲍勃·克拉特基特，因为鲍勃不打算在圣诞节工作。当斯克鲁奇的侄子顺道拜访，并邀请斯克鲁奇一起吃圣诞晚餐时，斯克鲁奇这样答道：

我还能怎么办……当我生活在这样一个充满傻瓜的世界？……对你来说圣诞节是什么？除了没有钱付账；除了发现自己老了一岁而不是多了一小时；除了平衡你的账簿，里面有近一年来你欠下的每一笔账目。如果我可以实现我的愿望……每一个说“圣诞快乐”的傻瓜都应该和自己的布丁一起煮沸，然后在他的胸前插根冬青树枝埋起来。

是不是自我欺骗？的确是。狄更斯后来让我们看到了年轻的斯克鲁奇，他的童年虽然艰难，但是他和姐姐的关系很亲密，和他的第一个商业导师关系也很好，尤其是在他失去一个女朋友以后。我们看到，他实际上是一个体面的人，现在他把情感隐藏在自己的贪婪中。平安夜的晚上，一些鬼魂出现在斯克鲁奇的床边。恐怖和破坏随之而来。到了圣诞节早晨，真正的斯克鲁奇出现。他一醒来，就打开窗户喊道：

我不知道该怎么办！……我像羽毛一样轻盈，我像天使一样快乐，我

像小学生一样开心。我像醉汉一样高兴地发狂。祝所有人圣诞快乐！祝全世界新年快乐！

最后，他跑到侄子家，在这里，和鲍勃·克拉特基特一家人共进的晚餐正要开始：

现在，我跟你讲，我的朋友……我不会再忍受这类事情。所以……我准备给你涨工资……圣诞快乐，鲍勃……圣诞更快乐，鲍勃，我的好朋友，要比每一年都快乐。我会给你涨工资，并努力帮助你在困苦中挣扎的家庭，今天下午我们再讨论你的事情，边喝圣诞葡萄酒边聊，鲍勃。给火添点燃料，然后在你说出另一个“我”之前再买一个煤斗，鲍勃·克拉特基特！

狄更斯很好地捕捉到自我意识和自我欺骗的语言。斯克鲁奇最后一次讲话使用“我”的频率是第一次讲话的一半。开始时，他的情绪化稍微小一些。你可以看出第一次对话更有距离感。除了更多使用差异动词外，如“应该”，他通过更多使用“你”这样的词来将注意力从自身转移。狄更斯的语言唯一不合规则的是第二次对话中过度使用了具体名称和冠词。如果狄更斯在1832年写作的时候就有LIWC，他就会知道斯克鲁奇在书的开始处需要使用更多冠词。

◎网络发表状态中的自我欺骗

父亲去世的学生和斯克鲁奇的语言都属于与否认或躲避情感相关的自我欺骗。那么过度自信的自我欺骗呢？回答这一问题有一些简单而有趣的方式。一种是想到常用短语，非常确定一件事和基本确定一件事时，我们使用的短语不同。例如，短语“毫无疑问……”只会被极其自信（也许有人会说过度自信）的人所使用。但是短语“有可能……”可能会被更加犹豫不决的说话者所使用。

当人们以这两种不同的短语开端时，他们会说些什么？通过搜索引擎查询（比如谷歌、雅虎或必应），很容易找到答案。下面是每个检索得到的一些陈述。

毫无疑问：

表面的活着只能经历表面的自己。

高清电视是潮流。

所有参加暑期图片取景比赛的人都做出了超乎标准的作品。

21点是个有输赢的游戏，你可以在赌场有优势。

上帝会原谅婚前性行为。上帝不会因为一个人犯错而减少对他的爱。

拳击短裤代表这个星球的终极舒适度和时尚。

冲击帽子行业的将会是拳击帽头饰！

花点时间去发现你身边的养兔爱好者是值得的。

任何听到或读到这个惊人故事的人都会加入到拯救我们国家和民族青年的战斗中。

看这些绝对肯定无疑问的句子是件有趣的事情。也许只有我这么想，但是我对其中一些抱有怀疑态度。真的是每一位参加暑期图片取景比赛的人都做出了超乎标准的作品吗？拳击帽头饰真的会摧毁帽子行业吗？

这些确定性陈述的语言同样引人注目。以“毫无疑问”开始的句子包含少量“我”，积极情绪词汇频率较高，句子简单且较笼统。现在与更谦逊的确定性陈述做对比，下面是以“可能”开始的句子。

可能：

你怀孕了。比预约时间提前20分钟到，完成注册。

我会被禁足，如果我爸妈发现他们车的情况。

在特定情况下，你之前的伴侣有资格获取你第二次婚姻伴侣的财务信息。

你收到这个调查的来源不止一个；如果是这样，我为重复的情况提前道歉。

谁会对这些陈述表示不赞同呢？这些陈述的语气慎重且合理。你还会注意到，这些语言正式性更弱，更个人化。例如，以“可能”开始的句子使用更多的人称代词，尤其是“我”。句子更复杂、更具体。评估具体性的一个好方法是查看时间、地点和动作的指代。最后，犹豫不决的短语缺少情感，尤其是积极情绪词汇的使用。这几乎像是过度自信的作者用乐观来弥补真理。

◎推荐信中的自我欺骗

大多数大学老师每年都会写大量的推荐信。这些信是本科生读研和找工作用的。除了写信，我们还会读来自申请学校或工作的人的推荐信。多年来，我和我的同事发现，很难对推荐信进行区分。它们读起来都很积极向上。有没有一种方法能够将真正积极的信和形式上积极的信进行分类？

我认为自己是一个诚实、坦率的人。每当我坐下来写推荐信时，重要的是我说的都是真实的，但是同时要给学生积极公正的评价。我注意到一件事，当我开始写信时，我逐渐开始看到学生身上越来越多积极的方面。当我写完时，我相信我写的关于这位同学的内容“毫无疑问”非常完美。哎。这听起来有点自我欺骗。和同事聊天，很多人也有同样的感觉。

有没有可能使用语言分析来辨别哪些信是真实的，哪些是自我欺骗性的？为了找到答案，我分析了大约200封自己写的推荐信。我会先回过头来评估我对每一个学生潜力的真实感受。这一评估通常是在写信完成几年以后，不会受到写作后自我欺骗余光的影响。然后，我把信中我使用的词汇和我的评估联系起来，出现了四个可靠的模式。

在为那些我真正觉得会做得很好的学生写推荐信时，我往往：

·说得更多，句子更长，词更大。 我越觉得他们会表现更好时，写得越多。信越长，体现出关于这个学生我可说的越多。我的句子也会更复杂。

·使用积极情绪词汇更少。 这令我感到震惊。越是觉得这个学生有潜力，我越不倾向于使用类似“卓越”“美好”“优秀” 和“幸福” 这样的词汇。当我回去看原始信件时，原因变得非常明显。对于那些非常适合读研或非常适合某份工作的人，我倾向于给出他们具体有成就的例子。对于那些没那么适合的人，我往往只是说他们精于自己所做的事情。

·提供更多细节信息。 我注意到，在更有说服力的信中，我会把更多的精力放在学生们之前做过的事情上，而不是仅仅谈论学生本身。对于那些有潜力的学生，我在描述他们的工作和对项目的贡献上花的篇幅很不成比例。我使用了更多指代时间、空间和动作的词汇。相对来说，我的句子很少使用“他” 或“她” 这样的代词，因为我的关注点是项目而不是人。

·很少关注潜在读信者。 回想一下，人们使用代词揭示他们的关注点。很明显，在为稍微弱一点的学生写推荐信时，我的思绪会从学生飘到阅读者身上。我会使用如“正如你从候选人的简历中看到的……”或“你一定会认为……”这样的句子。

语言分析表明，我的推荐信表达的内容比我想象的要多。我怀疑，我

和大多数认真写信的人一样，在试图定位一个学生时没有刻意说谎。然而，毫无疑问，存在一定程度的自我欺骗。调查结果指出我是如何下意识地精心写信以呈现每个学生更好一面的。作为自我欺骗的一部分，其他写信者也会使用同样的技巧吗？还没有这一问题的相关研究。我的感觉是，自我欺骗的表达方式有很多种。不过，不同类型的自我欺骗似乎有着相同的语言标记。

想象创建一个自动分析推荐信的系统是件易事。可能有很多公司会对这个产品感兴趣。这是一个既有趣的又令人不安的想法。为了使计算机系统有效工作，我们需要从每个作者那里收集许多推荐信。很有可能，计算机会和人一样或稍微比人做得好一些。真正的问题是，推荐信目前对学生未来的表现不具有诊断力。计算机程序也许可以帮助筛选，但危险是，公司高管们也许会把程序的结果看得太重。要记得，语言分析是概率性的，推荐信通常只是雇人决定的一小部分。

[1] 本书已由机械工业出版社出版。

谎言、骗子和无赖的语言：寻找语言测谎器

自我欺骗和彻底欺骗之间的界线不是完全清晰的。那位自我欺骗的学生声称父亲的死对他来说没有什么问题，他真的是为了骗自己的家人和朋友，让他们以为他比实际要坚强吗？自我欺骗的推荐信写作者试图影响读信者，是在骗人吗？如果自我欺骗留下语言痕迹，那么这个痕迹和人们故意欺骗他人时留下的语言痕迹是否相似呢？

语言测谎器会激起父母、伴侣、老师、买车的人、执法人员，以及其他所有人、世世代代人的兴趣。理论上，开发一个这样的程序是可能的。西格蒙德·弗洛伊德认为，人们的真实情感有时会流露出来，并且可以通过微妙的语言错误看出，如口误。口误最有可能发生的情况是，一个人在讨论一件事的同时在考虑另外一件事。

只要开始关注，你就会发现口误非常常见——在对话、邮件、在线聊天，甚至是正式演讲中出现。就在上周，一个朋友上班时给我发了下面这样一条信息，她的老板很专横："下班后有时间一起喝杯咖啡吗？我想和你撂挑。"她打算说的是"聊天"（quick talk）而不是"撂挑"（quit talk）。我们见面后，我开始了我们的对话："告诉我你为什么想要辞职。"我的朋友对我高超的心理洞察能力赞叹不已。

通过语言揭示欺骗的方式有很多种。1994年，南卡罗来纳州尤宁市的居民们得知他们的邻居苏珊·史密斯的车在乡村道路上被劫持时都吓坏了，她的孩子们还在车上。史密斯称，一个非裔美国男性逼她下车，然后疾驰而去。超过一周时间，史密斯在国家电视台上对劫持者发出最有力的呼唤，希望他送回她的孩子们。在她成为官方嫌疑人之前，史密斯对记者说："我的孩子曾需要我。而我现在却帮不了他们。"

这个案子中，联邦调查局特工格外关注这一陈述。为什么她会用过去式谈论自己的孩子们呢？人只有在已经知道自己的孩子死亡之后才会用过去式讨论他们。几天后，史密斯承认她将车子开进湖里淹死了她的孩子们。她的动机是摆脱掉孩子们，以便嫁给当地的一个商人，这个商人不想被组合的家庭拖累。

***

从表面上看，建立一个语言测谎器应该很简单。我们找到了大量的例子，并且对比了叙述事实的人和说谎的人使用的语言。问题在于，人们欺骗的方式多种多样。有些是故意的，有些不是。有些严重的谎言一旦被发现就会带来生命变化，有些不过是室内游戏。有些谎言需要精心编造故

事，而有些只是在真实故事的基础上添加或忽略一些重要的信息。谎言的最终目的也各不相同，可能包括获得地位、金钱或性，避免羞辱或惩罚，或只是为了好玩。也许最重要的是撒谎的情境。一个人在被警方审讯时，和同一个人在试图让一对老夫妇购买不存在的海滨房产时使用语言的方式很可能会不同。虽然语言测谎器的研究还处于早期阶段，但是一些令人兴奋的语言模式已经开始显现。

◎捕捉虚假故事

苏珊·史密斯失败的一个原因是孩子失踪的故事和她的语言使用不匹配。讲述虚假故事比你想象的可能要难。令人信服的谎言需要描述事件和感受，这些你没有直接经历过。从某种意义上说，你没有像真实故事那样“拥有”一个虚构的故事。同样，在构建虚假故事时，你需要有根据地去推测真实情况下可能会发生什么。

真实的和假想的创伤 20世纪90年代，为了观察描述假想的创伤对健康是否有益，来自纽约州立大学石溪分校的梅兰妮·格林伯格、亚瑟·斯通和卡米尔·沃特曼做了一个奇妙的、有创意的表达性写作研究。首批招募的作者大约70人，他们生活中都曾经历过重大创伤。其中一半的人到实验室记述自己经历的创伤。每一个创伤的实质用几句话来总结。紧接着，另一半人来到实验室，但是这组人不记述自己的创伤。相反，每个人拿到某个人已经写好的简单的创伤总结。这组人被要求想象他们经历过这一创伤，然后记述它，仿佛它真实在自己身上发生过一样。

所以，一半的文章是人们真实的创伤经历，另一半是想象出来的。两组文章的感染力都非常惊人。事实上，研究人员发现，记述想象中的创伤和真实创伤，两者疗效几乎一样。在阅读这些文章时，人们通常不可能知道哪些是真实的，哪些是编造的。

然而，通过计算机分析两组数据，发现了很多惊人的不同。首先，记述人们自己真实创伤的文章比记述想象创伤的文章要长。他们能够对自己做过什么、没做过什么提供更多的细节。其次，他们还使用更多的第一人称代词单数，如“我”和“我的”。再次，真实的故事比想象的故事使用更少的情感词汇，包括积极情绪和消极情绪词汇。最后，真实的故事使用更少的动词和认知（或思考）词汇。

这种混乱的结果并不那么令人惊奇。想一想每个语言差异的含意。记述真实的经历与下述特点相关。

·**词更多、词更大、数字更多、细节更多。** 如果在生活中真正经历过创伤，那么对你来说，描述它是件容易的事情。你不需要做很多思考就能

够描述这次经历的细节。一些细节信息包含时间、地点和动作。

·情绪词汇和认知词汇更少。 如果经历过创伤，你的情绪状态就会很明显。例如，如果你的父亲去世了，大多数人不会说“我真的很难过”。它隐藏在经历中。然而，没有经历过死亡的人会想：“嗯……如果我的父亲死了，我肯定会非常难过，我应该把它写在文章中。”同样，经历过创伤的人已经有一个合理的故事来解释它。这个正在编造故事的人肯定会做更多的思考，也就是使用更多的认知词汇去解释。

·动词更少。 有很多种动词可以起到不同的语言功能。当一个人使用动词时，通常表示他们正在提及一个更积极、更有活力的事件。对于一个经历过创伤的人，多数情况下这件事已经过去了。如果你正在描述一件假想的创伤，你在描述的同时仿佛正在经历它。另外，假想的创伤会让人们问自己：“可能发生了什么事？我会做何感想？”差异动词，如“应该”，在假想创伤中使用的频率很高。

·更多自我指代：“我”。 “我”字表示人们正在关注自己——他们的感受、他们的痛苦、他们自己。同样，第一人称暗示一种拥有感。不奇怪的是，人们写自己的创伤经历时对自己的感受更敏感，同时将创伤视为己有。

这些语言维度用来区分参与者写的真实故事和假想故事很有效果。通过一些精妙的数据（交互效度分析），我们预估，我们的词类可以精确地对74%的真实或假想文章进行分类。这是一个有趣的案例，因为其中没有欺骗。

真实故事和虚构故事：斯蒂芬·格拉斯的例子 1995年，《新共和国周刊》（New Republic）聘用了一位年轻有为的作者，斯蒂芬·格拉斯，他刚从宾夕法尼亚大学毕业。几个月内，他开始发表评论，并就广泛议题发表了有影响力的文章。1995年12月～1998年5月，格拉斯发表了41篇知名度高且通常文字很美的文章。他的同事和读者们常常对他的能力感到惊奇，他总能找到如此有趣的人采访，还能转述他们有趣的言论。

看来，格拉斯仅仅通过编造就能获取这些有趣的故事和言论。当他最终被杂志社的编辑发现后，他被开除了。一项调查显示，他的文章中至少有6篇是完全编造的，有21篇部分编造，剩下14篇可能是真实的。格拉斯的丑闻最终拍成了一部出色的电影《欲盖弥彰》（Shattered Glass）。讽刺的命运转折，格拉斯最后去了法学院，大概是去追寻真相与正义。

有没有可能格拉斯在编造的故事中留下了语言痕迹。使用计算机语言分析程序，我找出了它们。总的来说，语言分析结果显示，完全编造的故

事和可能真实的故事彼此有很大不同，部分编造的故事处于两者之间。格拉斯的真实故事和可能真实的故事使用的词类有以下特点。

·更多词汇、更多细节、更多数字

·情绪词汇（尤其是积极情绪词汇）和认知词汇较少

·动词较少

·自我指代较少："我"

听起来熟悉吗？除了自我指代，格拉斯真实故事使用的语言和人们记述自己创伤时使用的语言相似。他编造的故事和实验中记述假想创伤的语言痕迹几乎相同。

很有趣和很重要的一个例外是"我"的使用。上文提到，记述自己创伤的人使用"我"的频率比记述假想创伤的人多。这里的解释是，记述自己的创伤会引发更多真实情感，拥有感更强。在此基础上，斯蒂芬·格拉斯在编造故事时使用更多的"我"就很容易理解了。看一下他最后一个可能是真实故事的导语。

"试一试1、2、3、4"，亚历克·鲍德温说道，他清了清嗓子，"试一试1、2、3、4。"亚历克·鲍德温出演过《猎杀红色十月》（The Hunt for Red October）和《大亨游戏》（Glengarry Glen Ross），他把麦克风拿到离嘴十几厘米的地方，然后自豪地盯着它。"这个大巴车有麦克风。"他对我们几个说，我们聚在一起观看他初次亮相基层政治。

——《新共和国周刊》，1997年12月8日

对比他最后一个完全编造的故事的开端。

伊恩·瑞斯蒂是一个15岁的电脑黑客，他看起来像青少年版的比尔·盖茨，他正在发脾气。"我想要钱。我想要一辆马自达。我想去迪士尼乐园玩。我想要《X-战警》（X-Man）漫画书第一部。我想要终身订阅《花花公子》（Playboy），然后额外奉送《阁楼》（Penthouse）杂志。给我钱！给我钱！这个男孩穿着一件磨破的卡尔·瑞普肯的T恤，一遍又一遍地喊出自己的需求。"

——《新共和国周刊》，1998年5月18日

与真实故事相比，格拉斯编造的故事写作风格更浮夸。你可以感受到他在将不可能的事编造成为事实经历时的兴奋。你可以感受到他的骄傲和

对故事的拥有感——甚至是编造大胆与欺骗性故事的兴奋。在完全编造的故事中，格拉斯作为“公正”的作者和在虚构的引用中使用“我”的频率令人目不暇接。当撒谎时，克拉斯流露出认真的骄傲和自我关注，大多数人在讲述真事时才会这样。

◎捕捉虚假看法和信念

销售、政治职位候选人、管理人员和其他位居权威的人有时会公布信念，后来我们会发现并非如此。信奉虚假信念的艺术不止被政客或犯罪头目所掌握。我们中有多少人曾经持有我们不相信的看法，为的是讨好某个有吸引力的、强大的或可能有用的朋友？我们应该为自己感到羞耻。幸运的是，我们自我欺骗的能力完好无损，足以让我们知道真正的问题在于别人用他们欺骗的言论来欺骗我们。

有关一个热门话题的看法：堕胎和选择　人们在表达自己关于感情话题的真正信念时，计算机程序能够很好地检测到吗？

几年前，为了回答这个问题，我和马特·纽曼、黛安·贝瑞、简·理查兹进行了一系列的实验。我们招募了大约200名学生，要求他们对高度情感化的堕胎话题提供自己的两种观点：一个他们相信的观点和一个他们不相信的观点。一些学生被要求在家完成这两篇文章，然后晚些时候发邮件给我们。另一组学生只是在心理学系的实验室隔间里把两篇文章打出来。还有一部分学生被要求在录像时大声说出他们的真实信念和虚假信念。

如果你和大多数人一样的话，那你应该对堕胎这个问题有一个相对清晰的观点。有些人认为这是女性的选择，另外一些人反对堕胎的施行。想象一下你现在被要求写一篇支持自己观点的文章和一篇反对自己真实观点的文章。这可能是一件令人不快的任务，但是多数人可以完成它。

人们可以辨别出哪个是你的真实看法吗？我们招募了几个学生来阅读这400多篇文章，让他们猜哪篇是作者的真实观点，哪些不是。学生判断的正确率为52%，而猜对的概率是50%。换句话说，在堕胎问题上，让读者辨别人们真实观点是件困难的事情。

计算机的判断结果好很多，正确预测人们真实观点的比例为67%。诚实话语标记很大程度上与我们其他的一些研究相重叠。在表达真实观点时，学生使用的词更多、更大，并且使用的句子更长、更复杂。他们的论述更细致入微，较少情绪化。尤其有趣的是，他们使用专属词汇（例如，除……之外　）相对较多。当人们区分某事物属于某类或不属于某类时，会使用专属词汇；“我做了这个，没有做那个”。当人们使用专属词汇描述他们的信念时，他们试图在信念周围画线，以阐明他们相信什么和不相信什

么。

在叙述真实观点的时候，人们使用更多的自我指代，使用“我”的频率高很多。在叙述虚假的观点时，学生们表达了更多的积极情绪。

实验升级：“有偿”的态度　写出或谈论自己关于情感话题的真实观点和相对观点确切来讲不是关于欺骗的重大测试。有些人认为，这根本就不是欺骗的例子。几年前，保罗·艾克曼、莫林·奥沙利文和马克·弗兰克开发了一种更有说服力的方法。艾克曼一直被认为是第一代非语言交流研究的专家。除了描绘跨文化情绪图以外，他还研究人们被引诱撒谎时的面部表情。几年前，我听说艾克曼发表了一个引人入胜的演讲，这个演讲是关于测谎的近期研究的。他做了一个实验，实验中的20个人是科学和戏剧的完美结合。

想象你得知一个实验进行一小时左右，并且会付钱给你。你打电话报名，然后你在邮件中收到一份问卷，问你对近期一些话题的看法，比如死刑、抽烟、环境保护引动。几天后，有人打电话给你，预约参与研究的时间。你被告知，艾克曼教授会就其中一个话题对你进行简短的访谈，这个话题你持有坚定的信念。有些人被要求讲述他们的真实信念，另一些人讲述他们与问卷结果相反的虚假信念。你还得知，艾克曼会和你聊几分钟，然后试图判断你是否在表达真实的信念。

下面是有趣的部分：如果你告诉他真相，他相信你说的是真话，你会收到10美元的奖金。而如果你撒谎，他认为你说的是真话，你会收到50美元的奖金。然而，如果他认为你在撒谎，你将得不到任何奖励，事实上，你可能会到噪音室接受惩罚。噪音室是一间小黑屋，你必须独自在里面坐一个小时左右，时不时会听到巨大噪音。换句话说，努力让艾克曼相信你说的是真相，对你来说是有利的。

艾克曼的研究小组将访谈做成录像呈现给各种各样的人观看，包括心理学家、地方和国家执法人员、受过审讯培训的高级别联邦政府官员，并要求他们区分哪些人说的是真实信念，哪些人在撒谎。总的来说，正确率在51%～73%不等，猜对的概率同样是50%。

在听了艾克曼的演讲后，我问他是否愿意将访谈的文本分享给我，我可以用计算机程序来分析它们。我们的安排是：他给我们文本但是不告诉我们谁的话是真实的。然后我返回我的结论——有关谁说的是真实信念和谁在撒谎。几周后，我分析完所有数据并做出决断。艾克曼的合著者莫林·奥沙利文几乎立刻就给予回复，说我的工作非常了不起。在这次小量样本的基础上，计算机预测的正确率介于65%和75%之间。

艾克曼项目显示的与欺骗相关词汇组与其他研究发现的词汇组基本一致。也就是说，那些与艾克曼交谈时说实话的人使用的词更多、更大，使用的句子更长、更复杂，并且表达的积极情绪比说谎的人要少。还有，和之前一样，说真话的人使用更多的“我”。

犯“罪”后苦撑　艾克曼项目要求人们尝试在面对面的互动中欺骗某人。某种意义上讲，它是关于学生对某个特定话题看法的意志考验。这些学生没有做什么错事，也没有什么不良行为而被质疑基本的诚信。一种更加分明的研究欺骗的实验方法是真正诱导人们做出可疑的行为，然后在他们许可的情况下，对他们的所作所为向调查人员撒谎。完成这些的一个标准方法称为“模拟犯罪”。这个想法是，参与者同意“偷”东西（通常是钱），然后“被抓”时，他们向研究人员撒谎，研究人员并不知道他们是否偷了钱。

几年前，我们和马特·纽曼合作完成了这样的研究。签字同意参与实验的同学首先会见到马特，他会向大家解释，首先他们会被送到一间屋子待几分钟。一旦在屋子里坐下，他们就要看椅子旁边的书，并且翻到第160页。如果那一页有钱，他们要偷走钱，然后合上书。之后，他们被告知，有人会进入房间询问他们是否拿了钱。他们要否认拿钱的事情。所有人都接受规则。

进入房间后，一半的学生发现了钱（一张一美元的钞票），另一半没有钱。然后另一个实验员进来，翻看第160页，然后问：“这里没有钱，是不是你拿的？”所有人都回答没有。然后，实验员宣布，学生们会被带到另外一间屋子接受讯问，以确定他们说的是否是实话。讯问规模是相当小的，只要求学生详细说他们进屋之后做了什么。学生的陈述转录成文本后用计算机分析，和我们其他的项目一样，我们捕捉撒谎者的概率比单纯猜测要高很多。

模拟犯罪研究和“有偿”态度研究都发现相同的结果：语言中有可靠的“标识”可以为欺骗提供线索。此后不久，一些实验开始测试语言-欺骗间的联系。朱迪·伯贡是沟通领域最受尊敬的研究者之一，她进行的实验数字惊人，实验表明，不同类型的欺骗，尤其是自然互动中的欺骗，有自己的语言痕迹。她不止一次地证明，基于实验的欺骗研究适用于群组而不是大学生。在美国不同监狱的男性和女性罪犯的欺骗任务中，加里·邦德和他的同事们发现了相似的语言效果。

虽然这些研究令人印象深刻，但是关于各种欺骗项目的批评反复出现，那就是几乎所有的项目都是基于设计好的实验研究。事实上，多数研究和室内游戏惊人地相似。最糟糕的是，如果有参与者“被抓”，他们会失去很少的钱，可能和玩一把赌注较小的扑克牌差不多。如果赌注是真实的

且有可能改变命运的话，欺骗的语言标记又会是怎样的呢？

事关紧要时捕捉欺骗：避免入狱、心脏病和战争

控制性实验的优点是，研究人员可以清晰地看出什么导致什么。涉及生死的真实世界的实验却复杂很多。研究人员通常没有办法掌控局势，而且很难找到一个情境，在这个情境中你可以确定无疑哪些人在撒谎，哪些人在说实话。

◎证人席上的谎言：伪证和免罪

在公开我们的一些欺骗研究后，我收到了我见过的最有趣的研究生入学申请。申请人丹尼斯·赫德尔在过去的21年里成功经营着一家私人调查公司。她准备退休，觉得有必要到研究生院学习知识，然后建立一个基于语言分析的、安全的谎言检测系统。她申请书中的所有内容都指出，她是一个非常聪明并极为执着的人。我们很快见了面并一致同意：研究生院不适合她。相反，我会与她合作，来开发一个更现实世界版的语言测谎器。

丹尼斯的想法是：在现实世界中找到我和马特·纽曼进行的模拟犯罪研究的类似物。丹尼斯花费上千个小时在法庭上，目睹上百人在法庭上作证——很多人在撒谎。几周时间，我和她策划了一个富有想象力的研究。我们（指丹尼斯）会找到一些犯有重罪的人的法庭记录，这些人在证人席上明显在撒谎。美国司法体系的某些案例中，被判定犯有重罪的被告同时会被判定犯有伪证罪。伪证罪通常是有非常明确的证据证明被告在法庭上撒谎。（给未来罪犯的提醒：如果你是一名正在受审的罪犯，并且有非常有力的证据证明你犯罪，那么不要撒谎。你会因此感谢我的。）

我们同样需要一组人，他们明显没有撒谎但被判定有罪。幸好，丹尼斯找到了11个这样的人，他们被判定有罪，但是因为DNA或其他有力证据而被无罪释放。这是一个重要的对比组，因为无罪释放的人说的绝对是真话。

从我的角度看，这听起来肯定是一个简单的项目。只是找到这些录音，然后放到计算机里跑一跑，然后，砰！名誉和荣誉随之而来。正如经常发生的一样，这没有我想象的那么简单。

丹尼斯花了几乎一年的时间找到这11个无罪释放和35个犯重罪的案例。为了符合这次研究的要求，人们必须在证人席上作证过，而且录音必须完整。丹尼斯有一个小拖车，她会开到联邦法院的存档处。每天，她需要使用法院的复印机复印法庭记录，有时候会有几百页长。晚上她再返回到拖车内，将复印件扫描到她的计算机中。通常她会在拖车内一次待上两周，然后给自己放几天假。回到家后，她还会花额外的时间仔细阅读法庭

记录，挑选出对陪审团的决定起到关键作用的部分和那些无争议的部分。

丹尼斯的工作最终没有白费。虽然案例数量有些小，但是结果很有意义。最引人注目的是代词的使用差异。和其他多数研究一样，无罪的被告使用第一人称单数的频率比那些犯有重罪和伪证罪的被告高很多。“我”代表无辜。有趣的是，真正有罪的被告使用更多的第三人称。他们试图将罪责从自身转移给他人。另外，和很多早期研究一致，说真话的人使用的词更大，描述事件时提供更多细节，而且思考方式更复杂。

这些模式和结果的影响力是显著的。我非常激动，但是丹尼斯很失望。计算机对这些案例的正确分类率为76%，同样，猜对的概率是50%，这比陪审团的效果要好，但是比丹尼斯预期的95%差了很远。在撰写本书时，丹尼斯仍保持乐观地认为基于语言的谎言检测系统是可能实现的。也许她是对的，但是如果她当时可以得到识别欺骗的正确率达80%的可靠系统，那将会是宇宙性的突破。

检测欺骗：现实中的一课

我的职业在很多情况下与执法机关的相关部门有来往，我一直对有关欺骗检测的荒谬言论感到印象深刻。具有良好声誉的研究人员花费多年时间研究欺骗的生物学相关因素（类似于测谎仪测试的因素）、欺骗的非语言标志、欺骗的语言以及现在研究的欺骗的大脑活动，所有这些都走向同样的结论：我们有一个比靠猜测效果更好的系统。至今还没有任何一个捕捉说谎者的可靠度高于65%的系统。即使是那些命中率在这附近的系统，操作环境也是在高度控制和人为制造的情境下。

虽然如此，我曾不止一次地听说某个人或某个公司声称拥有命中率95%的欺骗检测系统。这不可能。这在我们有生之年是不可能实现的。

有趣的是，测谎仪的证据在法庭是不被允许的。测谎仪其实很令人印象深刻，因为它识别有罪人的正确率接近60%～65%。法庭允许的目击者识别的正确率可能和测谎仪不相上下。

是时候考虑在法庭使用科学的证据了。具体来说，所有都是概率性的。如果测谎仪、非语言行为、目击者、脑部扫描和其他任何类型的证据可以帮助我们对有罪和无罪证人进行的话，那么它们应该被引入法庭。然而，它们应该以为陪审团校准正确率的方式被引入。每种类型的证据都需要陪审团去衡量，都存在问题。生命是概率性的，法庭证据也一样。

◎为爱撒谎：评估潜在恋爱伙伴的诚实度

和决定某人后半生是否应该入狱相比，决定是否与某人约会没有那么庄重。抛开入狱的比喻，在线交友网站可以决定你应该和谁度过后半生。在挑选约会对象和可能的人生伴侣时，识别欺骗是重要的事情。有传言称，人们在网上描述的自己有时是虚假的。

不如问问杰夫·汉考克。汉考克和他在康奈尔大学的同事们做了一个吸引人的实验，实验对象是纽约市区域的80名网上交友者。这些人（一半男性一半女性）选自四个商业交友网站，基于他们在网上的简介。所有简介都包含照片、身高、体重、年龄、自我介绍和兴趣爱好。同意参与实验之后，每个人来到实验室，并在那里完成一些调查问卷。每个人都会被拍照，他们的驾驶证会被扫描，并且会测量他们的身高和体重。

正如你看到的，汉考克已经确认过他们的年龄（从驾驶证上获取）、身高和体重，他能够确定这些交友者的在线信息是否真实。他还可以找一组评价者，对比他们在实验室拍的照片和在网上展示的照片。男性倾向在身高上撒谎，女性倾向在体重上撒谎。不论男性还是女性在网上展示的照片都比在实验室拍的照片好看，虽然有些人的网上照片比其他人的还要夸张。通过收集的所有客观信息，汉考克为每个人计算了欺骗指数。

这次实验研究的核心是确定，最具欺骗性和交友者和诚实的交友者的线上个人信息在语言使用上是否有不同。确实，他们之间是有区别的，而且有别于其他的欺骗研究。那些较诚实的人倾向写得更多，使用更大的词和更长的句子，较少使用情绪词汇（尤其是积极情绪词汇）。诚实与否的最好预测，毫不吃惊，是“我”的使用。

虽然功能词区分了诚实和虚假的在线个人简介，但是在内容词上也有差别。那就是，对自己介绍不诚实的人倾向于将重点从敏感话上转移。例如，谎称自己体重的男性和女性不太可能谈及食物、餐馆或饮食。同样，那些照片具有欺骗性的人重点强调工作和成就来塑造自己的地位，对自己的外表则轻描淡写。

你真的需要计算机来评估在线推荐的人是否值得信赖吗？我们不能凭直觉和所有信息来挑选出诚实的人吗？汉考克的研究团队描绘了一幅惨淡的画面。他征询了50名学生的帮助，来评价每个在线简介的可靠性。这些学生在判断诚实度上没有比你扔硬币做决定的结果好多少。一个原因是，我们在猜测谁可靠时往往关注的正好是错误的语言线索。作为线上个人信息评价者，我们多数人认为乐观的、积极的、简单的、无私的和脚踏实地的人是诚实的人。而这些，我的朋友，就是为什么你的“爱情检测器”是有缺陷的仪器。

◎为战争欺骗：伊拉克战争之前领导人的语言

任何历史专业的学生都会回首1950年左右美国和伊拉克之间的关系，并且会问：“他们在想什么？”“他们”指两个国家。也许最让人疑惑的事件转折点是美国决定2003年3月入侵伊拉克。

2001年“9·11”事件，世界贸易中心和五角大楼被袭击后不久，乔治W.布什当局认为伊拉克可能参与了袭击。这里没有介绍美国和伊拉克之间错综复杂的历史，需要注意的是，两国已有嫌隙。之后的一年半，布什当局开始提高对伊拉克关注，包括声称伊拉克窝藏恐怖分子，制造大规模杀伤性武器，以及计划攻打西方国家。这些担忧的大部分是没有根据的，为的是事后利益。没有恐怖分子训练基地，没有大规模杀伤性武器，也没有攻击任何人的计划。尽管如此，有关伊拉克危险的说辞日益增加，这帮助推动了2003年3月入侵和占领伊拉克的行为。

发动战争的民主政府的社会动态是非常复杂的。在高度焦虑时期，谣言蔓延迅速，人们很容易受到歪曲信息侵害。传闻和猜测在不断重复的情况下，慢慢地转变为坚定的信念。欺骗和自我欺骗之间的区别可以很快地消退。

美国公共廉正中心（Center for Public Integrity，CPI）是一个独立的监察机构，一个非营利性和无党派组织，支持拥有广泛议题的调查性新闻（www.publicintegrity.org ）。在入侵伊拉克几年后，公共廉正中心开始梳理“9·11”事件和入侵伊拉克的战争期间由重要政府官员发表的有关伊拉克的公开言论。的确，任何人都可以获得上千篇声明，包括演讲、新闻发布会、公开的短片以及电视和广播上的访谈。对于研究人员来说这一任务更容易，因为那些已经证实是虚假内容的文本会被突出显示。这会使你认为……一个字数研究人员会只将突出显示部分和未突出显示部分进行对比。

事实上，研究在线交友欺骗的杰夫·汉考克分析了公共廉正中心的数据库。康奈尔团队整理了532篇声明，每篇至少含有一条客观上虚假的言论，同时含有对等数量相同来源的真实言论。这些声明发表于2001年9月11日和2003年9月11日之间，发表人是来自布什政府内部的八名高级人士：布什本人、副总统迪克·切尼、国务卿科林·鲍威尔、国防部部长唐纳德·拉姆斯菲尔德、国家安全顾问康多莉扎·赖斯、国防部副部长保罗·沃尔福威茨，以及白宫发言人阿里·弗莱舍和斯科特·麦克莱伦。

公共廉正中心数据库中的一个例子是副总统切尼于2002年3月24日在美国有线电视新闻网的《最新版本》（Late Edition）节目上的访谈。访问者沃尔夫·布利泽问切尼，他是否支持联合国派遣武器核查人员进入伊拉克，试图找到大规模杀伤性武器存在的证据。切尼回答（注意，突出显示的部分被公共廉正中心认为是虚假的，没有突出显示的部分被认为是真实

的）：

我们所说的是，沃尔夫，如果你倒回去看记录的话，问题不是核查人员。问题是他有化学武器并且使用过它们。问题是它正在研发并已经拥有生物武器。问题是他一直在推行核武器。是大规模杀伤性武器和他已经用它们所做的事情。本周的《纽约客》杂志刊登了一篇灾难性的故事，故事发生在1988年，他使用化学武器攻击了伊拉克北部的库尔德人；也许还袭击了200个不同的城镇和村庄。如果文章可信的话，据文章报道，遇难人员高达10万人。

这是一个十分邪恶的人，正如总统所说。而且，他此时正在积极推行核武器，我们认为这引起我们和该区域的每一个人的担忧。在我的旅途过程中，我发现这确实也引起了我们那里的朋友的巨大担忧。

计算机文本分析对比真实的和虚假的陈述，得出的结果和其他的研究具有可比性。真实的陈述与更高频率的“我”相关联。另外，它们更具有细微差别，重点在更多的细节，倾向于使用更少的情感词汇。上面的引用是展示这些区别的一个很好的例子。真实的部分中，切尼提供的信息更详细，使用更复杂的句子，使用“我”更多。

和其他的欺骗研究项目一样，这可能可以估计出计算机挑选真实和虚假陈述的效果如何。汉考克的团队称，计算机区分真实和虚假陈述的正确率为76%，同样猜对的比例是50%。

然而，解读这些数字时要谨慎。在我们知道谎言存在的时候，区分真话和谎言变得容易得多。如果试图猜测一个没有欺骗迹象的人陈述的可信度，那么我们的猜测很可能是无法解释的。在确定欺骗时，有两个交叉的问题：第一，我们不知道“地面实况”是什么样的。也就是说，我们不知道实际上哪部分是真实的；第二，我们不知道真实的基础比例是多少。如果我们知道切尼的从句有50%是虚假的，那么我们对计算机输出结果的解读方式和只有2%是虚假的情况会不同。在伊拉克战争开战前数月中，不存在可靠的地面实况，也肯定对欺骗的程度没有概念。另外，让问题变得更为麻烦的是，很多发表虚假陈述的政府人员很可能对他们自己的言论深信不疑。

监察：哪些词汇是真实和欺骗最好的预言者

这一章描述了几种类型的欺骗——从情绪否定和过度自信的自我欺骗到人们为避免牢狱之灾而进行的有意欺骗。即使不同类型欺骗的动机、策略和人都有很大不同，但是语言模式却惊人地相似。

当然，每一个项目都很特别。然而，随着研究都已落定，我们可以旁观，然后进行“倾向测试”。“倾向测试”是通过各种研究来评估是否存在可靠的欺骗模式的非科学性方法。正如你可以看到，下面的图6-1中有五类词类。每一类词右侧都对应有一个长条。长条距离页面右侧越近，说明该词类与诚实和讲真话更相关。长条远在左侧，说明人们在撒谎或欺骗时通常会使用这一词类，长条位于中间的词类与欺骗无关或无法进行分类。

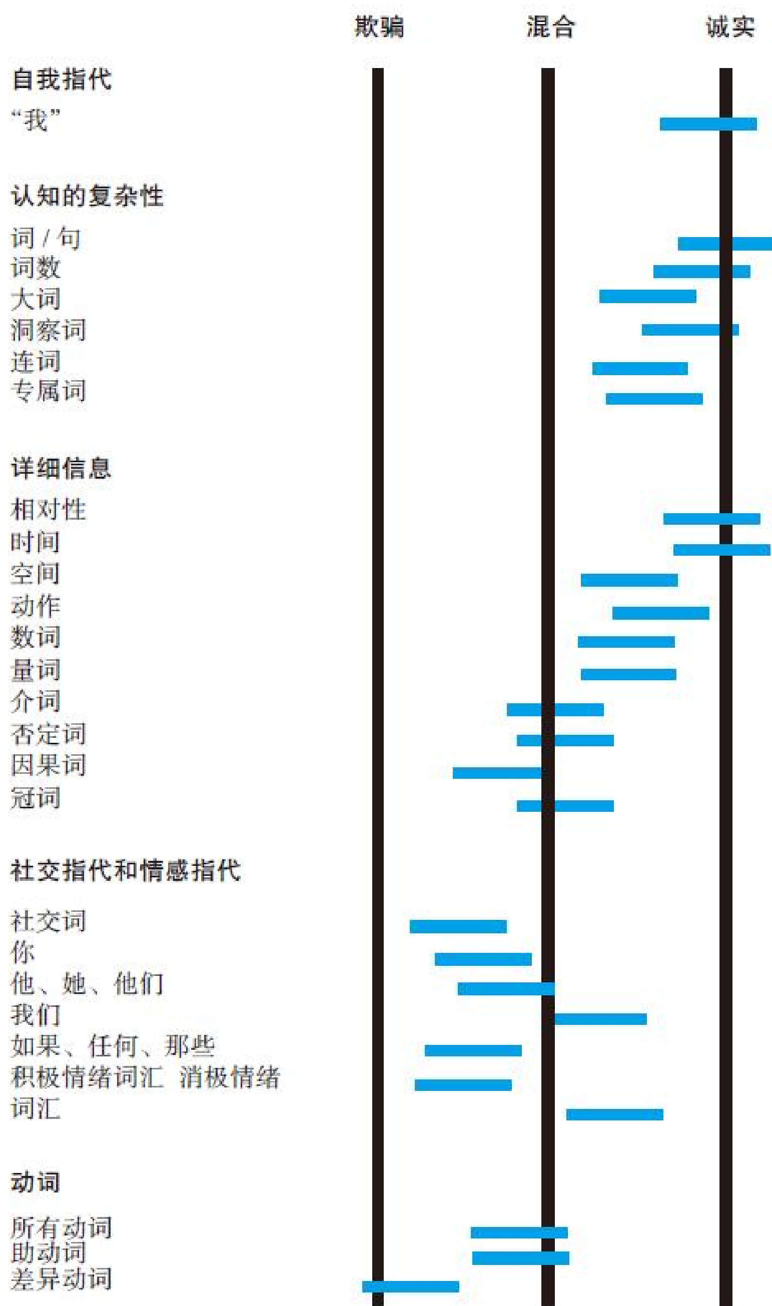


图6-1　诚实和欺骗相关词类

注：长条在左侧的词类与欺骗相关度更高。越靠右侧，可信度越高。右侧是诚实的标志。长条接近中点的词类与真实和欺骗的相关度都不高。

◎自我指代："我"

几乎在本书的每一章中，第一人称代词单数都非常的重要。欺骗研究中，"我"是一个人诚实的最好的独立标志。

"我"字的使用具有巨大的社交意义和心理学意义。根据定义，它是一种身份陈述。在对话中使用"我"是在向对话中的同伴表示你意识到你自己，你正在关注你自己。这样做在一定程度上有弱点，尤其是如果你的同伴正在评价你或者正在寻找某种方式来伤害你时。我常常认为"我"是一个微妙的顺从姿态，就像一只地位低下的狗翻过身将肚子暴露给另外一只危险的大狗一样。"嘿，我不是你的问题。我乐意为你效劳。我不会对你构成威胁。"

已经有一些研究表明，当人们不得不关注自己时，他们会变得更加谦逊和诚实。罗伯特·维克朗德来自挪威的卑尔根大学，20世纪70年代，他提出一个自我意识的理论。他和他的同事们想出几十个富有想象力的研究，研究中人们会在两种情况下做某种任务：在镜子前和不在镜子前。如果在镜子前完成一份调查问卷，他们会有较低的自尊，通常积极情绪较少。更有趣的是，他们的答案倾向于更真实，例如他们的身高、年级和行为和他们真实的身高、年级和行为更匹配。另外，在镜子前完成问卷会让人们使用更多的"我"。

为什么自我关注会使人们更诚实？维克朗德假定，关注自己会使人们短暂地沉思他们理想中想要成为的样子。也许他们终其一生的梦想是变得坚强、诚实、美丽、勇敢和热情。面对镜子使他们意识到自己没有成为理想中的样子。然后，最终人们会看到真实的自己和理想中的自己之间的鸿沟，这会使他们对自己感觉不太好，但是同时，也会激励他们努力成为更好的自己。在维克朗德看来，自我意识驱使我们所有人成为自己想要成为的样子。确实，我们大多数人通常想要对其他人诚实，而对自己诚实。自我关注激发诚实。"我"字只是反映自我关注。各样的研究表明，当我们看到"我"增多时，很有可能自我关注度增高。而且，自我关注的同时，人们倾向于更诚实。

◎认知的复杂性

通常情况下，人们讲述事实时产生的故事的复杂性高于虚假的故事。人们在讲述事实时不仅说得更多，而且使用的每一个句子都更长、更复杂。他们用的词更大，表明他们的陈述更精准、更具细微差别。讲真话者

的陈述通常更具有思考性，使用像是“意识到”“理解”“思考” 之类的洞察词汇。

讲真话者使用句子更长的一个原因是他们用连词将更多短语连接起来。连词包含的词，如“和”“或者”“但是”“由于” 等。很多连词属于专属词汇，如“但是”“或者”“除了……”“没有……”“排除……” 。正如前面章节所述，人们在作区分时会使用专属词汇。它们用来区分已发生的事和未发生的事，区分人们在考虑的事情和没有在考虑的事情，区分属于某分类的事物和不属于该分类的事物。

如果你是在撒谎，那么谈论你没有做过、没有见过或者没有思考过的事物是一件极其困难的任务。如果你是在讲述一件纯虚构的故事，那么所有你描述的事情都是自己没有经历过的。编造一个虚假故事，努力描述你没有做过的事情时会很头疼。大多数谎言通常由简单直接的陈述构成，陈述的是一个人可能做过或看到的事情。关于没做过的事情，人们很少有具体的评论。

◎详细信息

当描述一个真实事件时，我们关于这件事和事件发生时我们的具体情况有着丰富的记忆。更具体地讲，我们明确知道当时我们身体所处位置。我们的身体和相关事件存在于真实世界的三维空间。所以不奇怪，描述一件真正发生的事情时，我们会自然地描述时间、地点和动作。所有这些维度由相对词汇捕捉。从图6-1可以看出，讲真话者更有可能使用引发时间（例如“在……之前”“十点钟”“早晨” ）、空间（例如“在……上面”“下一个”“在……周围” ）和动作（例如“去过”“放置”“跳跃” ）的词。这些维度中，讲真话者在数词和量词使用上更具体，例如“更多”“更少”“几乎没有”“更大” 。

表示更高认知复杂性以及表示详细信息的词汇使用与近期有关目击证人审问过程中的陈述研究完美匹配。阿尔德特·维吉是分析嫌疑犯访谈的世界级专家之一，他指出，谎言的检测取决于人们发问的方式。一名警察、一位家长或朋友在控诉式访谈中很可能只是做出简短的、充满否认的陈述——这会变得非常难以评估。相反，维吉和他的同事们建议，审问和访谈应该更加开放，审判性质不那么强，把重点放在收集信息上。允许嫌疑犯说得越多，这样他们越可能证明无罪或有罪。

虽然维吉和其他法律执行研究人员近期才开始使用基于计算机的方法，但是他们已经独立发现，讲真话者的陈述更复杂、更详细。

◎社交和情感方面

很多方式的欺骗都与乐观和过度自信相关。努力卖给你一个新的毯子，向你推荐新的信仰或者推崇一场战争的人通常充满对真相的确定。高效的销售人员的一种方法就是使你相信一旦购买这个产品，毫无疑问，你就会和销售人员一样幸福、一样自信。

和常识相反，欺骗者更多地指代其他人，依赖更多的积极情感词汇。典型的骗子是寂寞的、鬼祟的、诡诈的、自我嫌恶的、奸诈的和紧张不安的人，他们拼命地躲避别人的视线。也许身边这样的骗子不多，但是我还是要力劝你，当你遇到多话的、热情的、乐于助人的、自信的人向你提供一个很好的交易时，请握紧你的钱包。

顺便说一句，社交词汇是表示社交关系的词的组合，包含的名词类似“朋友”“伙伴”和“母亲”，还包含像是说话、打电话和倾听这样的动作。阅读过很多欺骗文本，让我印象深刻的是人们为了证实自己的陈述或为了把责任转嫁到其他人身上而频繁地提及其他人。

◎动词和动作

动词很复杂。在一个语言专家（比如斯蒂芬·平克）周围闲荡几个小时，你的思绪就会遨游在各类动词的海洋中。规则动词通常表示某个特定动作，可区分出过去时和一般现在时。助动词，有时称为辅助动词，其实只是少数动词，例如“成为”（to be）、“拥有”（to have）和“去做”（to do）。助动词和被动语态相关联，美式英语课堂为之皱眉，英式英语课堂为之欢庆。另一种类型，差异动词（或情态动词），包含词汇“应该”“必须”和“将要”这样的词。人们在提出事实本是如何和事实可以如何或应该如何之间的某种微妙差异时会使用差异动词。

正如你在图6-1中看到的，人们使用动词频率越高，欺骗性越强。这一模式尤其适用于助动词和情态动词。假设你是一名小学老师，你的三个学生给出几乎一样的理由。

1.I finished my homework but the dog ate it.

我完成了我的作业，但是狗吃了它。

2.I had finished the homework but the dog must have eaten it.

我已经完成了作业，但是那只狗肯定已经吃了它。

3.The homework was finished but must have been eaten by the dog.

作业被完成，但是肯定被那只狗吃了。

第一个理由更有可能是真实的。它包含两个动词过去式，表明动作是具体的以及被完全完成。第二个理由使用了五个动词，暗示动作没有被完成，而且情态动词“肯定”（must）表示可能没有发生。第三个人的理由是三个中最低级的谎言：六个动词、使用过去时、没有一个“我”字。

英语中，动词提供大量关于动作的信息。它们暗示一个动作正在进行、部分完成或者全部完成。有些动词，例如情态动词，微妙地表达出一个动作可能已经发生，但是可能没发生。说那只狗“可能”或“肯定”已经吃了作业有力地暗示了那只狗的行为，但是同时拉远了说话人和其表达的实际发生行为之间的距离。

◎其他常见欺骗标志

使用情态动词是我们试图误导他人的几种方式之一，但是与此同时，从技术角度讲，又不属于撒谎。还有一些我特别喜欢的标志，如下所示。

被动结构：“被犯下的错误。”在一本愉快的关于误传的书《错不在我》[Mistake Were Mode（But Not By Me）]中，卡罗尔·塔夫里斯和埃略特·阿伦森研究人们如何通过巧妙地语言策略频繁地逃避责任。例如，历史学家普遍认为，20世纪70年代，国务卿亨利·基辛格频频欺骗美国人民越南战争的趋势和范围。几年以后，在一次采访中，塔夫里斯和阿伦森引用基辛格的话说道：“错误应该是我服务的政府所犯下的。”注意他的措辞。显然，基辛格没有犯任何错。相反，某人可能犯错了。

避免回答问题。在模拟犯罪实验中，学生被要求“偷走”一美元，我们直截了当地问每一个人：“你有没有偷走书中的一美元？”真正拿走钱的人会说一些如下这样的话。

我不主张偷窃。我有阴影。很久以前我曾偷过一次；我当时……还小。我很不喜欢明知他们要来抓我的那种感觉。我只是，你知道，尤其是你说为了一美元？我不会拿它的。

我为什么要拿？我甚至都不会想到要去这本书中寻找美元。我只是在写日记记录我的新生研讨课。

你们这样指责我，真的让我感到很不舒服。我永远不会做这样的事。

讲述事实的人最常见的回答是：“没有，我没有拿你的美元。”和撒谎者不同，讲真话者直接回答问题而没有任何修饰。正如这些例子所证明的，当有人没有直接回答你的问题时，他们很有可能在对你隐瞒什么，无论他们听起来多么认真。

让我明确表达：行为句。 语言学家和哲学家很长一段时间被称为行为句的语言策略所吸引。行为句是关于陈述的陈述。在陈述句“我向你保证我没有偷钱”中，短句“我向你保证”就是行为句。它只是在表达“我跟你说”或者“我要向你说出下面的话”。行为句有意思的地方是，没有办法估计它们的真实度。在由“我向你保证”开始的句子中，“我没有偷钱”没有被直接表达。这个句子的真实情况是说话人只是在说，他保证没有偷钱。这是一个细微差别，但是被使用的频率高得惊人。

任期结束之际，总统比尔·克林顿被记者追踪，问题有关与白宫助手莫妮卡·莱温斯基发生不端性行为的传闻。1998年1月26日，发布会上，克林顿宣布如下。

我将要再次说明：我没有和那个女性，莫妮卡·莱温斯基，发生性关系。

一个天真的人类可能会认为总统确实没有和莱温斯基发生性行为。实际上，他的表述是正确的：“我将要再次说明……”从技术角度讲，这是正确的。他是又说了一遍。好吧，虽然他后来承认曾与“那个女性”发生性关系，但是正式来讲，在新闻发布会上他没有撒谎。

罗杰·克莱门斯一直是一个优秀的棒球投手，他被之前的队友指控在棒球生涯中服用提高成绩的药物。实际上，在承认曾经服用过药物之前，克莱门斯在一次新闻发布会上说了如下一段话。

我想毫无保留地明确声明：在我的棒球生涯中，事实上，是我整个人生中的任何时刻，我都没有服用过类固醇、人类生长激素或其他任何被禁止的物质……

同样，从技术角度讲，你可以看到克莱门斯先生是诚实的。他的确，毫无保留地明确声明。他所声明的是一个谎言，但是确实是一个声明。

***

大体上，我们所有人都会向我们自己、向我们的朋友和整个世界撒谎。大多数这些谎言是为了避免伤害他人的感受或试图让我们自己看起来比原本更好一些而编造的无害的谎言。同样，我们也经常受到来自朋友、政客、广告客户和其他所有人的欺骗。当然，谎言并不全是坏事。它们使我们的生活变得更加有趣。正如电影《谎言的诞生》（The Invention of Lying）的演员，同时也是编剧和导演的里奇·格威斯所称，没有谎言，我们就没有小说，就没有好的故事。换句话说，没有谎言和欺骗，这个世界对艺术、文学、哲学和心理学的需求就会非常少。

欺骗他人的首要工具就是语言。毫不吃惊，通常谎言表现在我们使用的功能词上。正如其他章节所述，功能词本质上具有社交性。它们体现我们和物体、事件的关系，最重要的是，它们体现我们和其他人的关系。正如我们已经看到的，谎言最终反映说话者和听者之间关系的微妙转变。在欺骗发生的瞬间，人们的关系发生改变，这一切在功能词的使用上有所体现。

第7章　语言体现地位、权力和领导力

高地位的人在跟低地位的人讲话时，较少使用“我”和“我的”。相反，低地位的人往往使用“我”的频率较高。

这部电影是《教父》（The Godfather）。场景是纽约市的一间会议室。桌前坐着来自五大家族的黑手党头目，包括教父维托·柯里昂。会议讨论的是家族们是否应该沾手非法的毒品交易。从会议中获益最多的两个头目是菲利普·塔塔基利亚和埃米利奥·巴西尼。柯里昂召集了会议，他怀疑塔塔基利亚是谋杀他儿子的幕后黑手。经过一番讨论后，一项解决方案达成。

巴西尼：那么我们就一致同意了。允许毒品交易，但是需要被管控，同时，柯里昂会放弃对东部的保护，那里也会有秩序。

塔塔基利亚：但是我必须要柯里昂严格的保证，随着时间的过去，他的地位会越来越稳固，他会不会采取个人报复？

巴西尼：你看，我们都是讲理的人，我们没必要像律师似的给予保证……

下一个场景中，柯里昂坐着他的豪华轿车和他的律师汤姆·哈根离开了会议。哈根认为下一步他会和塔塔基利亚见面，讨论协议的细节。柯里昂打断他，指出真正做决定的人其实是巴西尼，而不是塔塔基利亚：“塔塔基利亚只是一个小混混……但是我直到今天才知道原来一直是巴西尼在背后搞鬼。”

哇哦。教父是怎么知道巴西尼才是真正做决定的那个人呢？在会议室场景中，巴西尼表现得泰然自若，而塔塔基利亚表现得更僵硬和紧张。两个人的语言也有所不同。巴西尼主要使用的代词是“我们”，而塔塔基利亚使用“我”。事实上，“我们”一直是高地位指示器，“我”表明更低地位。虽然《教父》的编剧马里奥·普佐并不知道代词研究中“我们”与权力和地位的相关性，但他直觉地知道如何塑造笔下人物的语言。

如果你混迹在一群谋杀犯中，发现真正首领的能力能救你的性命。事实上，了解任何团体由谁负责都显示出适应性。工作中，我们的上司可以提拔或开除我们。杂志编辑可以接收和拒绝我们的论文。一个暖人的微笑或眉毛挑起，高中啦啦队长或橄榄球四分卫可以使我们变得受欢迎或被嫌弃。作为社交性动物，我们深陷社会层级中。蚁群、狗群、黑猩猩群、小学操场、会议室和疗养院，我们都可以看到其中的社会层级。不论身处哪

个团体中，我们大多数人会努力融入，同时争取获得尽可能多的影响力。

关于社会地位的话题，存在一些内在的令人不愉快的感觉。清醒地意识到它，嗯，似乎不合适。总之，我们大多数人生来就被教育，所有人都是平等的。不幸的是，套用乔治·奥威尔的话，有些人比其他人更平等。严峻的事实是，要想成为一个健康的人，你一定要留心社会地位。确实，那些在生活中幸福和成功的人尤其擅长在社会层级中解码和工作。

弄清楚谁是头儿：解码社会层级

我们所有人都发现自己身处一群认识彼此的陌生人中。作为新成员的部分谜团就是努力弄清楚谁是谁，以及所有人如何融为一体。一些涉及理解社会层级的心理计算包含：弄清楚谁是负责人，谁是新人或谁在此情况下感到不自在。可以获得两种类型的信息：非言语行为和语言线索。

◎地位的非言语标志

研究人员试图确定处于高社会层级和低社会层级的人们是如何表现的。这不像看起来那么容易。经过对几十个很好的科学研究进行细致周到的分析，东北大学的朱迪思·霍尔只能够确定一小部分与地位相关联的行为。

·说话声音大　高地位的人往往比低地位的人说话更大声。

·打断　高地位的人比低地位的人更爱打断人。

·肢体亲密度　高地位的人比低地位的人更喜欢站得或坐得离人更近。

·放开倾向　高地位的人更具有放开身体的倾向，意思是他们的胳膊和腿都更伸展。低地位的人四肢往往位于比较闭合的位置。

在熟记这四点与地位相关联的非言语因素之前，你应该知道，即使是这些方面，也不是那么可靠。霍尔总结道，有些非言语因素在某些情况下可靠，但是在其他情况下却不一定。比如说，当两个处于同等社会地位的人交谈时，说话声音大这一点可以适度预测主导地位。然而，当两个处于不同社会地位的人交谈时，位于低社会地位的人往往说话声音更大。

霍尔的分析之所以如此重要，是因为多数人认为与地位相关联的行为其实与地位并不相关。例如，多数人认为高地位的人更少感到坐立不安、更少微笑、讲话更快、声音更放松和低沉、更常触摸他人、站得离人更远。不是这样的。换句话说，你所认为能够体现在一个群体中的地位和权力的行为可能根本无法体现。

人们不擅长察觉到一间屋子中比较有影响力的人的一个原因是，比较没有影响力的人通常会做一些事情，使自己显得比实际上更重要。几年前，我参加了一个测试人们识别地位能力的博物馆展览。视频展示了一系列简短的办公室相遇，相遇者是两两的成年男子。看视频的人来猜测两个

人中谁是老板，谁是下属。通常情况下，其中一个人坐得笔直，很少笑，身体稍向前倾。另一个人靠在椅背上坐着，看上去更随意。因为视频中没有声音，所以看视频的人只能通过非言语行为进行猜测。

大多数观看视频的人都错误地认为，那个严肃的、更僵硬的人是领导。实际上，视频内容是工作面试，坐得笔直的人是求职者。求职者努力让自己看上去有控制力、认真和不紧张。面试官只是在做自己。他们之前已经面试过几十个人，对自己的工作已经驾轻就熟。博物馆参观者上当了，因为他们认为，表现得处于主导地位的人就真的处于主导地位。

◎地位的语言标志

人们在对话、电子邮件和信中使用的语言能够很好地预测其在社会层级中的排位。正如你会猜到的，地位由功能词体现，而不是由所说的内容体现。在所有类型的功能词中，唯独能区分高地位与低地位说话者的词是代词。即使是在代词中，也只有一小部分词非常重要。

·低频率使用“我”　高地位的人比低地位的人使用第一人称单数的频率更低，例如“我”和“我的”。在任何两个人的互动中，高地位的人更少使用“我”。事实就是这样。高地位的人在跟低地位的人讲话时，较少使用“我”和“我的”。相反，低地位的人往往使用“我”的频率较高。

·高频率使用“我们”　高地位的人比低地位的人更常使用第一人称代词复数（“我们”和“我们的”）。

·高频率使用“你”　在书面和口语中，更常使用第二人称代词的人（“你”和“你们”）更可能处于较高地位。

根据情境的不同，其他类别的词汇也可能会与地位相关联。但实际上，代词“我”“我们”和“你”是至今为止一贯体现地位的词汇。在许多方面，代词与地位紧密关联并不那么令人吃惊。代词在所有词类中是社交性最强的，我们在对话中使用代词的频率极其高。第2章提到，代词反映人们的关注点。使用“我”频率高的人正在关注他们自己。那些使用“你”的人正在观察或考虑他们的听众。

事实上，有一些有趣的实验用来追踪人们在参与对话时的关注点。那些处于支配地位的人在讲话时往往关注他们的听众，在听讲时却会看向远方。低地位的人往往正好相反。他们在听讲时关注说话人，在讲话时却看向远方。他们在看什么？也许看向自己的内心深处。

“我”反映对自身的关注非常合理，与此同时，符合低社会地位的

人。“你”的结果也同样合理。你可以想象，使用“你”就如同说话时用手指着另外一个人。

你听到我说什么了吗？

我说你呢。

你最好留心些。

如果你再吐一次泡泡……

“我”与关注点、“你”与地位之间的联系很合理。但是“我们”呢？之前的章节提到，“我们”比较复杂。表面来看，“我们”听上去温暖且不明确，而且理论上应该有一种群体团结感。问题在于，在与他人的对话中，“我们”实际上至少有五种不同的含意。

·“我们”指“我和你”。 这种“我们”，每个人都想成为其中一员。我和你要一杯咖啡。我们两个人都享受咖啡。“我和你”的“我们”有一个公共认知，就是某个特定的人或团体以及我都是这个团体中的一分子。事实上，我们的确有一个共同的身份。然而请注意，这里有一个微妙的问题。这可能有点儿冒昧。可能我认为你和我是一个团体，但是你不这样认为。事实上，正是因为这个原因，日语和汉语中，说话人在日常对话中使用“我们”时非常谨慎。假设对方没有共享这一团体感，对他们使用“我们”是无礼的。

·“我们”指“我的朋友，不包含你”。 你刚从和高中朋友的野营旅行中回来，正在和你的同事聊这件事。随着故事的展开，你可能会说些“然后我们吃早餐”这样的话。通常我们在和人聊天时，谈及某些行动或经历是你与他人共享但不与你的故事听众共享，这时告知听故事的人是有必要的。这种情况下，使用“我们”是排他的，以此表达出这一紧密型“我们”不包含你。

·“我们”指“其实是你们”。 我个人的最爱。这是指说话人虽然使用“我们”，但是是在礼貌地要求或告知他人一些事情。在开始上课同学们依然在和别人聊天时，我会说：“我们能别聊天了吗？”或者昨天晚上在饭店，服务生问：“我们已经决定好晚餐点什么了吗？”

·“我们”指“我”。 有时指皇室的“我们”，这类“我们”用来分散责任和暗指来自可能不存在的其他人的支持。我曾经听到一个管理人员对员工说：“我们觉得你做的这个表格不够准确。”唯一知道这个表格的只有那个管理人员。

·“我们”指“几乎地球上的每一个人”。 这是政治家的最爱，这类“我们”是最模糊的一种。“我们需要一个更好的政府。”这类“我们”突出的原因是几乎不可能列举出“我们”指谁。

回过头来看这五类“我们”，只有“我和你”的“我们”才是真正和个人相关的，才真正能够巩固或认识到说话者和听话者之间的连接。其他四类“我们”为对话的参与者建立了屏障。并不令人感到吃惊的是，随着人们社会层级的提高，他们使用远距离形式的“我们”越来越频繁。

总之，非言语行为和语言在正式和非正式群体中都可以体现出相对低位。以一个外人的角度看一群人互动，他们的各种行为时不时地在体现自己。问题在于，这些线索可能出现得很短暂，然后消失不见。一个领导可能说话声音更大，肢体语言更多，某些片刻使用更多的“我们”和“你们”——也许刚好能够被会议中的其他人注意到。如果以前的研究具有指导意义，那么多数观察者能够识别群体中的地位等级，且正确度比纯猜测稍高一些。然而，如果使用计算机分析群体的语言，你的正确率会高很多。

通过通信发觉自己的地位

我们来进行一下对个人的讨论。相对于你的朋友，你处于什么地位？比起其他朋友，你会不会更害怕某些朋友呢？你真的想知道吗？了解人际关系中地位的最好方式就是查看你和他人的往来。

这很简单。看一下最近十封你发给别人的邮件，然后和最近十封别人发给你的邮件做对比。计算出你们各自使用“我”的比例。如果你有很充足的时间，你也可以计算出“我们”和“你们”的使用频率。数据显示，“我”的可信度最高。规律如下：使用“我”更少的人社会地位更高。如果两个人使用“我”的频率基本相同，那么这两个人可能关系平等。

我知道这个，是因为我和马特·戴维斯几年前做过一个研究。我们招募了15名志愿者，有研究生、本科生和甚至还有几个教员，他们同意让我们分析他们和大概15个人来往的邮件。马特写了一个计算机程序打乱所有邮件，目的是让它们难以阅读，难以辨识他们真正说了什么，并保持匿名性。

除了邮件本身，我们让四名志愿者从几个方面评价和他们通信的那15个人：性别、年龄、了解他们的程度、喜欢他们的程度，以及类似的问题。关键问题是相对地位。也就是，对于这15个通信者，志愿者要在回答问题时做出判断。

就相对地位评价这位通信者：

1 2 3	4 5	6 7
另一个人具有	另一个人具有	另一个人具有
相对较低地位	相同地位	相对较高地位

要根据问卷上的七点评价所有的通信者。所以，如果志愿者是研究生，他们会给资深教授评6或7，给高中生评1或2。

结果是明确的。具有较高地位的人使用更少的“我”，更多的“你们”和“我们”。这不是一个细微的发现。对于大多数人来说，影响是相当大的。

思考调查结果时，我不停地问自己，我是否做了同样的事情。你应该知道，我一直都觉得自己是一个非常主张平等主义的人。我一直都努力怀有敬意地对待研究生、员工、社区居民和学术界中我的上级。无疑，在我的电子邮件中，这些分析就不会显示出大的地位影响。我是个平等主义者，记得吗？我分析了自己的数据，就在那一天，我放弃了我的“对待每个人都一样”的自我认识。对我的语言进行分析，得出了和其他研究相同的地位差异。下面是一些编辑过的例子。

敬爱的彭尼贝克博士：

我 上学期上过您的心理学导论。我 很喜欢您的课，并且我 从中学习到了很多。我 收到一封您的邮件，是关于和您做研究的事情。是否有时间让我 拜访您，讨论一下这件事呢？

帕姆

亲爱的帕姆：

这很好。这周不行，因为有旅行。下周二9：00~10：30

怎么样。见到你一定很高兴。

詹姆斯·彭尼贝克

你可能已经猜到，帕姆曾是我的心理学导论课上的大一学生。请注意她在每句中是如何使用“我”的。彭尼贝克回信中成功地避开使用单独的“我”。并非巧合，帕姆使用敬称“彭尼贝克博士”，彭尼贝克直接称呼她的名字。但是，现在看一下彭尼贝克在努力争取让一位世界闻名的教授参与会议时写的信。

尊敬的（著名的教授）：

我 写信的原因是我 在帮助组织一项关于（某特定话题）的会议……我 已经联系了很多人，他们都问您是否会参加。如果您能来的话，我 会非常高兴……唯一不足是，我 们无法支付任何费用……我 觉得最好的方式是把这当作一次重聚，而不是一次会议……我 真切地希望您能来。

詹姆斯·彭尼贝克

亲爱的詹姆斯：

很高兴收到你的来信。祝贺会议的举办。重聚的想法很好……会议能

为我们提供一个半正式的环境来跟进彼此近期的研究……有没有什么办法能让大学出几千块钱支付一下会议的路费？

祝好

著名的教授

突然间，之前那个冷淡的、高地位的彭尼贝克变成了低地位、有点卑躬屈膝的学生。如果看得更仔细一些，你会发现所有的邮件都很友好和热情。不使用“我”不会让作者显得冷漠或自大，只是会不那么平易近人和有一些距离感。

很多人思考这些调查结果，认为“我”字的使用差别只在做某种要求时产生。如果你需要某事物，你可能会使用“我”。处在施与方时，可能你就不会使用“我”。其他研究显示，情况并非如此。几年前我发出的一些电子邮件是一个不错的例子，当时我是我们部门的管理员。我们的办公空间不足，我不得不问几个人让他们搬到别的办公室去。下面是我发给一个非常重要的教授和一个谦逊的研究生的编辑过的邮件版本。结果在意料之中。

敬爱的（非常重要的教授）：

我 一直在努力避免这件事，但是我 想我 可能需要问您，是否愿意放弃您的办公室……我 可以为您找到只比现在稍微逊色一点点的地方。

亲爱的（谦逊的研究生）：

你可能知道，系里的办公空间现在短缺。我们在做最大的努力调解所有的学生。然而，你是否愿意搬移你的办公室呢……？非常感谢你的配合。

和之前的例子一样，邮件的语气非常诚恳、温暖且有建设性。虽然邮件内容基本相同，但是它们传达的信息却稍有不同。写给非常重要的教授的时，仿佛我的手中拿着帽子。你可以看到我微微弯下腰，轻声地在说话。在第二封邮件中，我根本没有在说话。我只是在传达“我们”的消息——来自系里。“我们真的不想打扰你，但是我们希望你能搬走。不是针对个人，请注意。”

我有意识地在调整这些邮件的语言吗？没有。我不记得有特别地去写它们。我的大脑自动运转，然后根据要阅读的人而调整了写作风格。不要怪我。是我的大脑对地位差异太敏感了。

我喜欢邮件研究，是因为它展示了一些重要的结果，更重要的是，揭

示了我自己的虚伪。地位层级无处不在，而且就在我们的眼前。通常我们会不注意它们，尤其是在和我们认识多年的好友相处时。

陌生人对话中社会层级的快速显现

回过头想一下这一章开始的《教父》情节。我们在自然会话中能够察觉到代词的使用吗，像维托·柯里昂一样，我们能够在任何场合判断谁是有影响力的人吗？一句话，不能。人们说话很快，我们凭耳朵无法察觉到代词使用的细微差别。然而，如果将对话转录成文字，计算机程序会遇到一点小问题。

和邮件或信件不同，当人们在和他人说话时很少深思熟虑，文字就那么从嘴里飞出来。忽略社会互动形式的不同，对话中可以识别社会地位的词类和书面沟通是一样的。有些与大学生的相关研究，另外不妨说，有些研究有关白宫高级政府官员。

在一个早期的与代词地位相关联的研究中，我的研究生埃娃·卡斯维兹邀请学生们参加“开始了解你”任务。成对的陌生人进入一个房间，讨论任何他们感兴趣的事情，时长10分钟，并会被录像。对话相对常规：你来自哪里？你的专业是什么？感觉大学怎么样？对话结束后，两个人进入单独的房间，各自评估这次互动。两个人往往就谁更有地位和谁更能控制谈话上意见一致。

埃娃项目的结果和其他类似项目的结果一样明确且一致。高地位的人使用“我们”和“你们”的频率高，使用“我”的频率低。事实上，同样的模式在人们上网聊天时也会出现。然而，最引人注目的是，对话中社会层级出现的速度很快。埃娃的第一个研究中，两个人一进实验室就开始了互动。他们被要求在坐好和录像机开机之前不要说话。

在对话开始一分钟内，社会层级就已经确定。我猜测，两个人在看到彼此的几秒钟内，对社会层级有一个粗略的感觉。也许他们注意到彼此的身高、体重、吸引力、穿着或整体举止。不管是什么，他们相对的代词使用几乎立刻固定。

两个陌生人在网上相遇聊天时，你可能会认为层级出现需要长一点的时间。令人惊奇的是，不用很长的时间。网上项目是我和同事山姆·戈斯林以及我的研究生学生依拉·陶斯兹克一起实施的课堂练习的一部分。上百个学生被要求在特定的时间上网，然后随机和另外一个人连接。我们只是要求他们聊任何自己感兴趣的事情，持续15分钟。15分钟结束后，他们会做一份有关这次对话的调查问卷。

正如面对面的交谈项目一样，地位的差异在互动的前三分钟就会显现出来。这令我们很吃惊，因为两个参与者无法得知对方的外貌、声音或其

他任何可以做出预先判断的信息。两个人怎么可以心照不宣地快速判断出谁更具有主导性。

阅读一些网上聊天的对话有助于理解这一过程。谈话初期，双方都会投掷出有关自己的信息，与此同时，也在搜取对方的信息。有些参与者还会做出微妙的举动来削弱对方的主导地位。下面是两位女士谈话的开始部分。女性B（布列塔尼）的话用粗体标出，对话结束后的问卷中，双方都认为她是占据主导地位的人。

A：嗨，有人吗？

B：有。嗨。

A：噢，嘿。

B：我是布列塔尼。

A：我们应该谈些什么呢？我是克里丝，很高兴认识你。

B：不要误会，你是男生还是女生？

A：哈哈。女生。

B：只从克里丝判断不出来，哈哈哈哈，很高兴认识你。你是哪个年级的？

A：大一。你呢？

B：酷。我大三。专业呢？

A：工作室艺术，但是我准备转去通信专业做新闻摄影工作。你呢？

B：历史。

A：你有在什么机构里工作吗？

B：嗯，我要拿到教高中的教师资格证了，但是没有其他的了。你呢？

A：我在考虑做一名摄影师。我不想一辈子画画。

B：摄影很好啊。

A：我不是一个好的艺术系学生……我高中学艺术是为了考进这里的艺术学校。

请注意在一开始时，两个人似乎以同样的方式在互动。然而，在布列塔尼写第三句话时，她使用了一种低调的高压攻势："不要误会，你是男生还是女生？"我问你，亲爱的，你可以想出任何一句以"不要误会……"开头的、让人听着舒服的话吗？布列塔尼已经让克里丝意识到，她有可能是一个会欺凌弱小的人。

之后，两个人都在寻找对方的信息，这些信息可以帮助她们确定所谓的社会层级。然而，当布列塔尼发现克里丝比自己小、专业没有自己好的时候，她从心理上占据了主导地位。当布列塔尼确立了自己的主导地位后，克里丝开始关注自己，并且更为紧张和顺从。后面的对话（在这里没有显示）中，较高地位的布列塔尼给克里丝提出一些关于如何可以（应该）获得教师资格证的建议。

有趣的是，克里丝和布列塔尼都很享受谈话过程并且很喜欢对方。在时间接近尾声的时候，两个人交换了邮箱地址并希望后续能够保持联系。可以肯定地说，如果她们约着喝咖啡什么的，布列塔尼仍然会是两个人中占据较高地位的人。

阅读成百上千个这样对话的过程中，观察社会层级显露非常有趣。"无知的"地位相关问题包括问另外一个人他们父母的职业、上个假期、运动习惯、种族划分、校园住宿位置、心理课成绩等。毫不吃惊，在某一领域地位较高的人通常是发问的那个人。请注意布列塔尼是如何问对方年纪和专业的事情的。

评估彼此的地位不是说英语的大学生的专利。这件事发生在世界的各个地方。事实上，有些群体评估地位的规则更简单。比如说在韩国，相对年龄是决定社会层级的一个非常严格的因素。假设年纪相同，那么接下来评估的是财富或收入。在这些社会群体中，人们直接询问别人的生活是很正常的现象，而在西方，这可能会被视为不礼貌。在最近一次去韩国的旅途中，我坐在一个和我年纪相仿的韩国男性旁边。在飞机飞行十分钟内，他开始询问我的年龄。当他清楚地了解到我们俩年纪一样时，他开始问我的年收入。当我回答完后，他温柔地微笑着说："咱俩挣得差不多。"我觉得他的收入应该远高于我，某种程度上，这使得他在后来的谈话中更自然。

定位尼克松当局领导人的地位

以实验的方式研究地位问题是对自然的社会过程的苍白模仿，通常需要数月或数年的时间去完成。现实世界中，成为老板的人很可能多年来一直被培养，并获得了组内其他人的尊重。有时会有捕捉知名人士自然谈话的机会，并且是在知晓地位级别的情况下。当这种情况发生时，我们可以开始观察这个团体的整体社会层级。研究相对地位一个很好的天然实验出现在一个时代前的录音带上。录音作为著名水门事件的一部分被公开，它惊动了整个美国并导致一位总统辞职。

理查德·米尔豪斯·尼克松1968年竞选成为总统，两年后，他在自己位于白宫的办公室里秘密安装了一个录音系统。只有个别职员知道这件事，从录音中可以听出，尼克松自己几乎不受隐藏话筒存在的限制。任职四年后，尼克松再次参与竞选，参与竞选的民主党候选人选票从一开始就落后于尼克松。部分原因是，尼克松政府手段猖獗，有些高层助手批准了一些非法活动，以帮助尼克松连任。一个阴谋是夜间闯入位于华盛顿水门办公大楼的民主党总部。这些夜贼在白宫工作，他们在竞选战略家的电话上安装了窃听器。由于一位守夜人的警觉，这些夜贼在离开大楼时被当地警方捕获。

因为这一盗窃行为太过奇特，所以很少有人能想到它由白宫责任人启动。闯入行为被发现的最初一个月，几乎没有人把它当回事，而且它的发生在接下来的四个月没有对竞选产生可觉察的影响。然而，来自《华盛顿邮报》的年轻记者鲍勃·伍德沃德和卡尔·伯恩斯坦坚持不懈地追踪新闻，直到尼克松最亲近的助理开始受到牵连。从1973年年初开始，水门事件成为头条新闻，直到尼克松1974年8月辞职。

调查的转折点是，中层白宫职员亚历山大·巴特菲尔德透露，尼克松在场时，白宫录音系统会录下谈话内容。接下来的数月，随着法律问题增多，白宫公开了一些转录的录音。当最后一段录音文字（也被称为“证据”磁带）在1974年夏天公之于众后，尼克松被迫辞职。

其中最切题的录音带涉及尼克松与他的律师约翰·迪恩、首席国内顾问约翰·埃利希曼和白宫办公厅主任哈利·罗宾斯·海德曼的谈话。这些公布于1974年春天的录音，对历史学家和像我这样的语言分析研究人员来说就是无主珍宝。最初的转录文字中，有15段尼克松与迪恩、埃利希曼和海德曼一对一的对话，我们能够分析这些对话。通过每一段对话，我们对比了尼克松和他的这些助理讲话的语言。

与实验研究一致，处于高地位的尼克松使用“我”的频率远小于他的助理们。总的来说，尼克松的语言中第一人称单数占3.9%，他的助理们使用第一人称单数的频率为5.4%。尼克松使用第一人称复数（1.4%和0.8%）和第二人称复数（3.4%和1.8%）的频率比他的助理们高。

进一步观察发现，这些文字显示出尼克松和这三个男性有着不同的关系。尼克松在与迪恩和埃利希曼的对话中使用第一人称单数的方式明显不同。这些代词模式表明，尼克松与迪恩和埃利希曼距离感较远，但是与海德曼的对话相对平等（见图7-1）。这是一个伟大的假设，但是否成立呢？

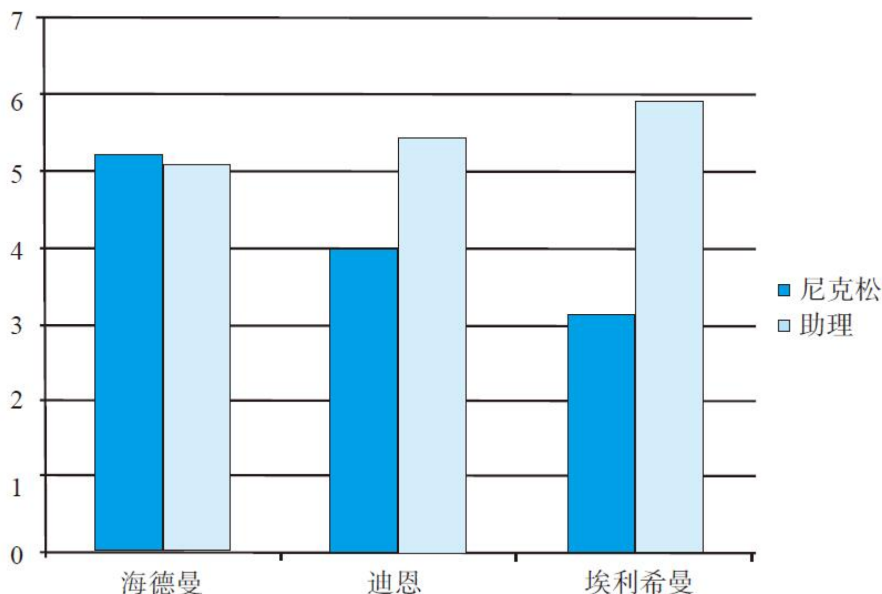


图7-1　尼克松和三个助理使用第一人称单数的频率

21世纪初期，当我们开始分析水门事件录音文字时，这四个人里只有约翰·迪恩还在世。有关尼克松和这三位助理的关系，他同意接受我的邮件采访。海德曼和尼克松早在20世纪50年代就认识彼此，用迪恩的话说，当时他们在政府管理方面势均力敌。“但是他们不是朋友。海德曼曾说过，尼克松不知道自己到底有多少个孩子，他们只握过一次手，就是尼克松开除他的那天。”但终究来讲，“这是一个合作伙伴关系。”

迪恩自己和尼克松的关系比较正式且彼此尊重。有趣的是，华盛顿政治超出了埃利希曼的能力，迪恩这样描述他：傲慢且缺乏安全感。在听水

门事件的录音时，迪恩对埃利希曼为得到海德曼的工作而上演权力的游戏感到印象深刻。埃利希曼在和尼克松的对话中过于殷勤，几乎是卑躬屈膝。尼克松的反应比对迪恩的态度更加疏远，和埃利希曼的关系更加正式。

对“我”字的最终分析是值得注意的。水门事件录音于1974年公开，录音录于1972年6月（闯入后不久）至1973年7月，这期间这件丑闻几乎每天都是头版新闻。请注意，享有更高地位和自尊心强的人倾向于过度自信，通常使用“我”的频率较低。1972年6月到12月，尼克松和自己助理们的对话中，使用“我”的频率处在2%～4%。而随着丑闻不断升级，尼克松的地位开始下降，他使用“我”的频率也逐月增多。1973年7月的最后一段录音，在和助理们的对话中，他使用“我”的频率徘徊在7%～8%。换句话说，随着尼克松的政治世界开始崩塌，他在和别人的交往中变得不那么占主导地位且不那么有影响力。

水门事件的录音文本显示出我们通过尼克松自己的眼睛（或者，也许是嘴巴）辨别尼克松助理们社会层级的能力。我们可以推断出，和其他人相比，如埃利希曼，尼克松对海德曼所说的话响应度和接受度最高。有趣的是，代词分析中无法得知助理们之间的相对社会层级。也许相对于埃利希曼，在总统办公室中，尼克松授予迪恩更高的地位。然而，当迪恩和埃利希曼同处于其他场景中，埃利希曼的地位高于迪恩也是可能的。只有通过对不同团体的分析，我们才能够逐渐揭开团体地位层级的复杂面纱。

最后，地位和喜好有很大的不同。正如迪恩指出，没有证据表明尼克松和海德曼很喜欢对方。尊重一个人和喜欢一个人是不一样的。

领导阶层的语言

对地位的研究可以延伸至对领导层的理解吗？有没有可能通过领导们的语言识别谁是优秀的领导？如果领导们改变语言方式，他们可以变得效率更高吗？令人吃惊的是，几乎没有研究分析过领导者的语言。更少有人为预测人们未来可能成为哪种类型的领导而研究人们日常使用的语言。

◎假设领导角色会影响领导的语言

确实，在小团体中，领导的说话方式不同于跟随者。伊桑·伯里斯来自德克萨斯大学麦库姆斯商学院，他和同事们进行了一个项目，大约有40组商学院学生完成同一项团体任务。每四个人为一个工作组，他们代表一家小型咨询公司，为一家虚构的公司提供咨询，以改善其客户服务部门。这项任务复杂，要求每个小组组员共同找出一个解决方案。

伯里斯项目脱颖而出，因为他指定了谁是领导者。他告诉每个组，基于每个学生性格测试的分数和分析，他已经选出最佳领导者。事实上，领导者的分配是随机的，类似于抽签。然而，重要的是，组员真的相信他们的领导者是有技能的。

每个组的互动会被录音及转录。结果与我们之前发现的结果一致。那些被指派角色的人使用“我”最少，使用“你们”和“我们”最多。换句话说，人们一旦拥有某种角色，语言会发生改变。 我无法强调这一发现的重要性。几乎所有关于领导者和语言的实验都是基于已经是领导或已经获得较高地位的人。这一研究表明，语言反映出领导角色。换句话说，如果人们成为具有较高地位的领导者后，大多数人开始说话像个领导。

◎是否可以由一个人的语言预测他的未来领导能力

每年在工业、政府、军队和其他大型组织为预测和选择未来的领导者而花费数亿美元。大多数专门从事领导力选择的咨询公司和他们所咨询的公司都是凭直觉和/或问卷的研究方法做出决定的。即使是最有科学依据的方法，在选择领导方面命中率也很低。

可靠地选择领导者是非常困难的。最大的问题是，一个人可以在某种环境中是一个出色的领导，但在另一种环境中非常差劲。我曾咨询过一家公司，他们公司的领导没有能力带领管理团队朝着让所有人满意的方向努力。同一个人，两年前在一家完全不同的公司里是一位非常出色的领导。他的方法、智力或者能力都没有改变。唯一改变的是环境。

第二个咨询公司遇到的问题在于建立评估方法。例如，一些公司会给未来的领导者大量的问卷。其他公司会让人经历折磨人的面试或案例分析。还有一些公司只是把他们未来的领导者安排到一家公司，让他们做演讲，认识一些人，然后参加一系列的会议。从两个相似背景的候选人中做选择，这些方法中的每一种都只是比用扔硬币来决定要好一点而已。

如果知道高效领导者的特定语言模式，我们是否可以通过自然语言来选择未来的领导？事实上，现在就做判断还为时尚早。语言分析可以帮助我们确定人们的思考方式、与他人联系的方式和看待自己的方式。预测人们在一个完全新奇的环境中，和一群具有不同动机且背景的捉摸不透的人在一起会做何反应，语言分析就没那么有用。

◎主导性的语言风格会使一个人成为更优秀的领导者吗

经过近百年的团队研究，心理学、社会学和商业领域的科学家现在明白，不同的团队根据任务、团队结构和团队中的人员而有效地工作。例如伯里斯项目中，领导者使用高地位语言越多，团队想出客观上更优质方案的可能性越大。但是，效果不那么大。另一个项目中，团队结构相对散漫，出现完全相反的结果：领导者的主导性语言导致团队表现更差。

我们回到了同样的问题，领导者的特点和需要完成的工作之间的紧张局势。许多研究表明，如果一个公司、一个体育团队或整个国家整体是有生产力的，少有暗斗，且在没有面临任何重大威胁的情况下，热情友好的领导者特别有效。这样的领导者可能不会表现出极其主导性的语言。但是如果组织面临着一系列的内忧外患，大多数人希望能有一个高效干练强大的领导者来完成工作。高效干练的领导风格特点通常包含清晰的指挥链和公认的社会地位。

◎领导者可以通过改变语言变得更高效吗

我认为，可以。但理由不是大家想的那样。只是使用不同的词汇不能自动地改变说话者的心理状态。正如本书所说，语言反映了我们的个性和社会情境，但很少直接影响它们。

在演讲、广告和简短互动中，个人使用的词语可能影响听众对他们的看法。事实上，如果说话者使用“我”或“我们”，我们的大脑的确会记录并相应地做出反应。另外，如果目标只是使一个人听上去像领导，那仔细修饰一下他们的语言在短期内会很有效。

但是，撰写演讲稿的人请注意。你希望你的总统候选人听上去像个总统还是接地气？听起来具有高地位还是低地位？两次总统竞选活动中，乔

治·布什比他的对手阿尔·戈尔或约翰·克里听起来更加个人化、接地气和低地位。2008年的选举中，巴拉克·奥巴马的语言比他的对手约翰·麦凯恩更加高地位和总统化。显然，候选人的语言会影响我们的看法，最终会影响我们的选票。

注意，获得领导地位与成为有效的领导者是不同的。多数情况下，高效领导者和其他人一起工作，和他们互动，做出大家都认同的决定。领导力不光做演讲。领导力需要大量的思考和社交技巧。

领导者可以通过首先倾听和分析他们自己的语言变得更高效。回想一下第2章描述的约翰·克里在2004年和乔治W.布什的总统竞选中使用“我们”的案例。他的顾问错误地认为，如果克里使用更多的“我们”，会给听众留下温暖和平易近人的印象。

如果我是克里的语言顾问，我会首先分析他的演讲和访谈。克里使用“我们”和“你们”的频率较高，使用“我”的频率较低，很快就能发现他在努力让自己听起来更总统化，给人傲慢的印象。我的语言策略是，努力改变他和听众的关系，以及他对自己的看法。显然，他需要知道，他使用“我们”的方式会让人觉得冰冷和虚假。

语言标志着和他人的关系，人们可以通过语言的使用使自己成为更优秀的领导者。语言就像车辆的速度计，反映车辆的行驶速度。你无法直接影响速度计来减速。相反，速度计是用来衡量你的驾驶速度的。将一个人训练为更好的司机，他们的速度计自然会跟上来。让潜在的领导者更加注意他们的语言及其含意，可以改变他们与他人的关系，成为更好的领导。

有关地位、权力、傲慢和领导力的最后思考

这种地位层次几乎是每一种关系的一部分，在一定程度上，有点儿令人沮丧和感到威胁。不仅是在互动的几分钟内就建立起社会等级，而且互动中的所有同伴都会采取他们的语言角色。坐在飞机上的人，聚会上遇到的陌生人，或者在商务会议中被介绍给同事，都会触发每个人去对比自己所处的社会层级。

这种几乎隐性的地位竞争在我们很小的时候就在起作用，并会一直持续到我们年老。虽然社会等级的存在可能违背我们对民主和平等的信念，但是我们建立社会等级快速有效的方式有助于使所有后续的互动更加顺利地进行。

虽然本章节中使用的例子主要来自英文项目，但是几乎相同的语言模式在古希腊语、阿拉伯语、日语和所有我们研究的欧洲语言中都很明显。例如，公元前五世纪，欧里庇得斯笔下的人物使用低频率的“我”和高频率的“我们”来彰显其强大。同样的人物，当失去权力后，使用“我”的频率飙升，而使用“我们”和“你们”的频率随之骤减。

你会注意到，人称代词（我、我们 和你 ）与地位、权力、自信心、傲慢和领导力相关。并不是说所有这些概念都一样。组织心理学家最近的研究表明，人们在没有直接压过他人权力的情况下也可获得高地位。高效领导力有时涉及傲慢和权力，有时不涉及。其他词语类别也许能够区分权力地位，两者都可以被认为是社会层级的标志。

这意味着，如果你未来在鸡尾酒会派对上遇到了甘地、斯大林、爱因斯坦和《教父》里的埃米里奥·巴尔西尼，你使用的“我”肯定比他们任何一个人都多。

第8章　爱的语言

那些自然地将他们的功能词彼此同步的情侣更可能在一段时间后仍保持他们的关系。

希望你现在已经相信，功能词反映心理状态。难过的是，我对你有所保留：语言使用不会发生在真空中。多数情况下，在我们和另一个人说话或打字的同时，他们也在和我们说话或打字。多数语言使用发生在人们持续不断的社交关系中。这意味着，我们可以使用语言工具来研究独立于个体以外的东西——我们可以开始研究人类关系。

考虑下这意味着什么。你和你的爱人、你的朋友、你的老板、你的敌人相互交谈的方式为你们关系的性质提供了重要线索。不只是你说话的方式——你和朋友发的邮件，你放在脸书上的帖子或发的微博，告诉了我们你整个的社交网络。

社会关系有自己的特点。我的社交网络中有一位朋友，我经常和他谈论个人话题或情感方面的话题，每次我们见面都去吃墨西哥餐；和另外一位朋友见面，我们总是喝酒、斗嘴和开玩笑。甚至是我和他们两个的邮件都能反映出我们周期性聚餐时的互动模式。这些不同的关系都持续了多年，并且都很温暖和令人满意。如果我来分析这些互动中的语言，就会发现它们都有自己的纹路，有自己的会话特征。

通过检查亲密关系的语言，关系本身的画面开始浮现。我们不仅可以看到数月或几年的关系，也可以追踪到短时间内两个人的连接方式。比如，可以通过分析两个人的相似程度来分辨两个人之间的相互关系。

看一下，下面是两个相爱的大学生的短信互动片段。

她：我很高兴至少现在可以跟你说话，但是我特别特别想见你。我讨厌跟你分开。

他：☺我希望这里出现一个迷人的微笑符号……我爱你。

她：说真的……我一直都是那种女孩，可以很长时间都不见也不会……想你，距离对我来说根本不是问题……但是今天，见不到你我觉得我都要疯了！就像我说的，我疯了似的想见你……疯了似的！！！！我也爱你。

他：我真的……很爱你。

这对情侣都密切关注彼此，并重复许多相同的单词和短语。对比一下表达强烈的爱的词汇和强烈的……嗯……愤怒的词汇。几年来，日间电视谈话节目The View在美国博得大量观众。主持人是一群活跃在新闻界、喜剧届和政界的聪明且固执己见的知名女性。多年来，偶尔会爆发激烈的分歧，产生的分歧既个人化又政治化。2007年5月23日，保守派伊丽莎白·哈塞尔贝克和罗茜·欧唐纳之间酝酿已久的紧张关系终于爆发。

哈塞尔贝克： 因为你是个成年人，所以我肯定不会是那个帮你解释你的想法的人。那是你的想法。你自己解释其中暗含的意思。

欧唐纳： 我为我的想法辩护。

哈塞尔贝克： 为你自己的想法辩护。

欧唐纳： 对，但是每次我为它们辩护时，伊丽莎白，我挑中的都是可怜的小伊丽莎白。

哈塞尔贝克： 你知道吗？可怜的小伊丽莎白不是可怜的小伊丽莎白。

欧唐纳： 对。这就是为什么我不会再和你争吵，因为这很荒诞。所以三个星期，你可以说任何你想说的代表共和党的废话。

哈塞尔贝克： 和唐纳德·特朗普这样的人争论更容易一些，不是吗？因为他令人感到不愉快。

欧唐纳： 我从没和他争论过。他和我吵过。我说了一件他的事情——

……

哈塞尔贝克： 我给你一个澄清的机会。

欧唐纳： 你没有给我任何东西。你也没有必要给我。我问了你一个问题。

哈塞尔贝克： 是我问了你一个问题。

欧唐纳： 你甚至都没有回答它。

哈塞尔贝克： 你甚至都不回答自己的问题。

欧唐纳： 噢，伊丽莎白，我不想——你知道吗？你真的不知道我在说什么？

哈塞尔贝克： 我知道你在说什么。

很明显，那对情侣之间互动的语气和这两个成年电视主持人是不同的。和那对情侣的短信不同，这两个女性的争吵类似于两个欺凌者之间的校园战斗。尽管两种互动存在显著差异，但是观察两组交流中在词汇使用上是如何融合的，这很有趣。两组交流中，两个人完全关注对方，几乎在模仿另外一个人的话：一组以爱的方式，另一组以喷射愤怒的方式。

对话就像跳舞。两个人毫不费力地跟随着对方的脚步，每个人通常都能预测到对方的下一个舞步。如果其中一个舞者转换至意料之外的方向，那么另一个通常会跟着转变并且形成新的步伐。就舞蹈来说，很难判断谁在领舞、谁在跟随，因为两个人在不断地互相影响。一旦舞蹈开始，几乎不可能单独一个人决定双方的动作。

威廉·巴特勒·叶芝在他的诗《在学童中间》（Among School Children）里问："我们怎能区分舞蹈与跳舞人？"通过阅读两个转录的文字，你发现你做不到。你几乎可以听出来参与者在调整他们说话的速度、语气和音调。随着他们变得越来越情绪化，他们的语言发生聚合。然而，稍微不那么明显的是，他们使用的功能词的相似性。开心的情侣和愤怒的脱口秀主持人都倾向于使用代词、介词、冠词和其他功能词，两个人使用的频率几乎相同。两组对话中，对话双方处于同样的心理状态中，并体现在他们的语言中。

语言模仿

社会科学家早就知道，面对面交谈中的人往往会表现出类似的非语言行为。当一个人交叉双腿时，另一个人会跟着做同样的动作。当一个人做了一个夸张的手势后，另一个人稍后也会这么做。

这种动作模仿最初被认为反映了双方很喜欢彼此。事实上，它是一个参与的标志，或表示两个人互相关注的程度。如果你正在恋爱或对你的对话伙伴感到愤怒，你们两个的非言语行为会互相匹配。

这也同样适用于对话中使用的语言。如果两个人在讨论同一个话题，那么他们的语言会相似。总之，这就是对话。更有趣的是，人们的说话方式也会聚合，即他们倾向采用同等水平的正式性、情绪性和认知复杂性。换句话说，人们倾向使用同类功能词，且频率相似。另外，两个人的关系越亲密，使用功能的匹配度越高。

功能词的匹配程度被称为语言风格匹配度（language style matching，LSM）。语言分析发现，语言风格匹配度发生在任何互动开始的前15～30秒，通常是无意识发生。几项研究表明，语言风格匹配在一些不太可能的地方很明显。

比如，想象一下，你被要求回答一系列开放式问题，这些问题是课堂作业的一部分。同样想象一下，每个开放式问题的写作风格都不同，从非常正式到非常不正式。你会注意到问题的风格不同吗？更重要的是，你会通过调整答案来改变你的回答方式来匹配问题的风格吗？惊人的是，你可能不会注意到问题的风格不同，但是你会调整答案的风格。

我和我的同事山姆·高斯林以及我的学生莫莉·艾尔兰，我们一起为几百个人的普通心理学课程设置了一个在线课堂写作作业。我们预先告知学生们，他们需要回答四个简短的关于课堂写作作业方面的问答题。我们没有告诉学生，不同问题使用了不同的写作风格。

例如，一个特定问题可能是浮夸的、傲慢的风格，或者另一个版本中，同一个问题可能是聊天式的“山谷女郎”的风格。这里有两个样本，要求学生写一个在教科书中讨论过的特定理论。

浮夸的版本： 虽然你的教授给了这个话题相当少的注意，但是认知失调是一个常见的心理现象，绝大多数完全不懂的外行人都熟悉这个理论……一旦一个人已经变得相当熟悉这个概念，举个例子应该很简单。

聊天式的版本：好的，我们可能很少讨论认知失调。我觉得非常疯狂，是因为就好像所有人都应该能够明白认知失调非常重要。就好像认知失调一直在发生，你知道我的意思吗？……所以，好了，到你了。我的意思是，认真想一个认知失调的例子，然后详细讲述给我。

在两个问题的最后，每个人都会阅读到这样一段话。

在下面空白处写出认知失调在现实生活中的例子，解释导致它的原因以及要如何解决这个问题。从书中寻找支持你例子的论据。

所以，你了解到两个问题的不同写作风格了吧。有趣的是，无论写作风格如何，学生们都给出同样有知识含量的答案。唯一的不同是，阅读到浮夸版本问题的学生，答案也是浮夸的风格，而那些阅读到聊天式版本问题的学生，答案中使用的也是同样自由风格的非正式术语。每个人需要回答四个问答题，每个问题都以不同的方式呈现，很多学生后来称，他们甚至没有注意到风格有所不同。

我们很难看着这两个问题，而不注意到它们措辞的显著差异。虽然如此，我们总是自然地调整以适应说话者（或出题者）。事实上，莫莉进行了另一个实验，她给人们两页以前出版的小说。然后，她让参与者接着原著作者停止的地方写出下一页小说。她明确告诉其中一半参与者要尝试匹配作者的写作风格。莫莉发现，所有人，即使是那些没有被告知要匹配风格的人也很自然地匹配了原作者的风格。事实上，当人们被直接告知要匹配风格后，匹配的结果会稍微差一些。

语言风格匹配可能比你想象的更普遍。你可能有这样的经历，看过一部特别吸引人的电影后，你说话就会像刚看到的某个电影人物一样。有人告诉我，在阅读一本具有独特写作风格的书之后，他们发现自己在接下来的几个小时甚至几天里会使用同样的风格写作和说话。事实上，如果我，喜欢，开始——你知道的——用“山谷女郎”方式写几个段落，然后，你知道，如果，嗯，你的电话响了，假设你接起了电话。你会非常想要这样说话。

我现在停止这样说话，以保持我们各自的尊严感。

语言风格匹配和大脑

如果风格匹配如此普遍，那么它为什么会发生呢？一种解释是，它是我们大脑中固有的存在模式。20世纪80年代，一个意大利神经科学家团队观察了猕猴的一组脑细胞的活动，只要猴子抓住一个特定的对象，这组细胞就会活跃。后来他们发现，当猴子看着一个人的手以同样的方式抓住同一个对象时，同一组细胞也会活跃。

其他研究表明，可能有一整套的脑细胞反映他人的行动。这些细胞群被统称为镜像神经元或镜像神经元系统。

最近的研究要求芭蕾舞演员在被扫描大脑活动时观看芭蕾舞视频。研究人员发现，观看芭蕾舞视频时，芭蕾舞者的镜神经元系统被激活，而非舞者没有显示相同的活动。我们大多数人都有过主观上的类似经历。例如，如果你曾经打过网球，你可能已经注意到，假如你在电视上观看激烈的比赛，你有时会注意到自己会稍微移动手臂试图打球。

特别值得注意的是，大多数研究人员报告说，镜像神经元在大脑中最密集的区域是布罗卡区。第2章中提到，布罗卡区也是涉及功能词处理的区域。使用功能词的能力与模仿非语言行为密切相关，并且根据最近的发现，与语音语调的变化也密切相关，这并不是偶然。此外，现在许多学者认为，模仿社会行为的能力及其与布罗卡区域的密切联系解释了语言能力的早期发展和进化。

镜像神经元研究现在依赖于最先进的脑成像方法，例如功能性磁共振成像，这些方法可以看出在非常具体的任务中，大脑哪些部分是活跃的。例如，普林斯顿大学的一个研究小组最近发现，在观看人们做情绪表情的短视频时，极具同情心的人的布罗卡区域比同情心较弱的人相比显示出更多的大脑活动。需要强调的是，镜像神经研究仍处于起步阶段，研究人员常在解释与同情相关的大脑活动的最佳方式上意见不一。

脑成像研究中出现的一个问题是镜像神经元和语言风格匹配之间的可能联系。

如何衡量语言风格匹配度

研究人员已经开发了多种方法来研究两个以相近频率使用功能词的人的匹配程度。一些技术的复杂程度令人难以置信。其他的方法还算简单。下面这种方法，事实上，你可以用计算器解决。基本想法是，我们想要找出任何两个文本在使用功能词上的相似程度。

虽然功能词有九类（包括人称代词、非人称代词、介词、冠词、连词、否定词、量词、常见副词和辅助动词），但人称代词在日常语言中最常见。虽然我们的计算机程序分析所有类型的功能词来计算LSM，但是仅通过观察人称代词的使用，你就可以对人们之间的互动有一个很好的概念。

回过头看一下本章开头罗丝·欧唐纳和伊丽莎白·哈塞尔贝克互相厌恶的对话。为了了解两个女性同步的程度，我们需要计算每个女性使用人称代词的比率，公式如下所示。

$$1-\frac{|\text{人物 1 代词比率}-\text{人物 2 代词比率}|}{\text{人物 1 代词比率}+\text{人物 2 代词比率}}$$

在她们简短的交流中，哈赛贝尔克使用了16个人称代词，总词数159；而欧唐纳使用的比例为24：182。换句话说，哈赛贝尔克使用人称代词的比例为10.1%，欧唐纳为13.2%。

竖线指绝对值。也就是说，人物1的人称代词比例和人物2的人称代词比例相减得到的值总为正值。所以，数值代入公式：

$$1-\frac{|10.1\%-13.2\%|}{10.1\%+13.2\%}$$

$$=1-(3.1/23.3)=1-0.13=0.87$$

瞧！只基于人称代词的LSM值高达0.87。有趣的是，当所有九类功能词都计算在内时，两个女性的LSM甚至更高，几乎完全一致。LSM值的范围为0～1，1表示两个人使用功能词的方式完美一致，0表示两个人使用功能词的方式完全不同。实际上，如果LSM值小于0.60，则两个人使用功能词的一致性较低，如果高于0.85，则一致性较高。

在这一点上，我们没有进行直接实验。但是，让学生回答开放式问题的风格匹配研究提到了一些类似的重要问题。那些很好回答不同写作风格问题的学生往往是在选择题考试中取得较高成绩的人。换句话说，那些认真听课的人会自然地更关注问答题。更关注的结果就是回答时在功能词的使用上与问题更一致。

我们都有镜像神经元，并有基本的模仿和同情他人的能力。这些能力

因人而异，也取决于我们在与谁交谈。我们都经历过这样的情境：我们不得不跟别人说些他们不感兴趣的话。承认吧，我们也经历过对方说的话我们不感兴趣的情境。我可能有一个充满镜像神经元的大脑，但如果飞机上我坐在一个陌生人旁边，听他描述自己的关节疼痛、药物历史和痰的观察，那我们使用功能词的方式肯定各自不同。

体验对话中的“LSM过山车”

镜像神经元帮助我们快速与他人的对话同步。从交流的角度来看，这是很有意义的。无论是与老朋友、商务伙伴、飞机上的陌生人，还是销售人员交谈，我们都会调整自己的语言风格，以建立一个共同的社会框架并减少摩擦。风格匹配有助于确保人们的情感基调、正式性和开放性相似，有助于理解他们彼此的相对地位。

此外，我们的大脑非常注意互动过程中语言风格的变化，不断纠正我们使用的方式。风格匹配在谈话的过程中不断波动。大多数对话中，风格匹配通常起点很高，然后随着人们继续交谈而不断下降。这种模式发生的原因是，在谈话的开始阶段，重要的是和另一个人连接。两个人都需要知道对方的想法和感受。随着对话的进行，说话者开始感到越来越舒服，他们的注意力开始转移。然而有时候，风格匹配度会立即升高。最好的例子是当会话主题或语调出现意外变化时。

想象一下，你正在为一个亲密的朋友计划一个惊喜的生日派对。在派对前一天，你的朋友提到她正在考虑第二天晚上去看电影，也就是你安排派对的那个晚上。你必须快速想出一个谎话，确保她会参加派对。从她的角度看，这是一个无关的对话；但是从你的角度看，这个对话充满感情。如果我们可以跟踪你们两个人的语言，那么能发现你们的风格匹配发生什么变化？令人惊讶的是，匹配度会上升，但不是你想的那样。

◎和说谎者对话

一旦一个人开始说谎，对话的语调和方向就会急剧变化。研究欺骗语言的康奈尔研究员杰夫·汉考克进行了一些实验来证明这一点。在一个项目中，他把成对的学生带到他的实验室，告诉他们需要在四个不同的话题上发生四次在线对话。换句话说，他们在不同房间的计算机上聊天，因此他们实际上无法看到对方。所以，两个合作伙伴都知道这一流程是如何运作的，他们首先需要以初识的方式聊天约五分钟。当两个人都对在线聊天系统感到舒服时，实验员会立刻给两个合作伙伴一个正式话题的列表，每个话题讨论五分钟。研究的狡猾部分是，两个合作伙伴中的一个收到了话题列表，对于其中两个话题的观点，他们需要明目张胆地说谎。他们的合作伙伴并不知道这一点，以为所有对话是诚实和坦率的。

我们已经知道，当人们撒谎时，他们的语言发生变化。你也许会认为，当有人撒谎时，语言风格匹配度会降低。事实上，在有人撒谎的期间，合作伙伴的语言风格匹配度也会升高。是的，说谎者说谎时语言确实

会发生改变，但是清白的讲真话者的语言也发生了改变，甚至改变得更多。

想想在这种情况下发生了什么。研究开始时，清白的合作伙伴已经与说谎者至少有过一段诚实的对话，然后突然间，另一个人的说话方式开始变得不同。我们的大脑高度关注改变。在这种情况下，清白的人察觉到有些东西“消失”，或者有些东西不合理。因此，他开始更加注意，以出现更高的风格匹配度。这显然不是一个有意识的过程，因为大多数清白的伙伴后来称，他们认为欺骗性对话看上去很正常。

关于这项研究，有一些问题总是让我感到困惑。回想一下惊喜生日派对的困境。如果我不得不欺骗我的朋友以说服她在第二天晚上待在家里，我必须快速思考，并想出一个合理的故事。说谎是一项艰苦的工作，撒谎期间，我不会给予朋友太多关注。相反，她会关注我。汉考克的发现表明，我朋友的语言使用将会适应我的欺骗性表达方式。

关于LSM有一个有趣的悖论。在撒谎的情况下，说话者之一（说谎的人）实际上较少注意对方。你会直观地认为这将导致风格匹配度降低。显然，与说谎者谈话时，我们开始花费更多的注意力来试图解码他们奇怪的语言变化。在说谎的情况下，一个人变得注意力下降，另一个变得注意力增强。令人吃惊的是，这发生的频率远超你的想象。

◎与分心的多任务处理者交谈

最近的研究表明，多任务处理非常常见且惊人的无效。当人们进行多任务时，他们试图同时做几件事情，净效果是他们所有任务的工作质量降低。我们都有过这样的经历，与人交谈的同时还在阅读短信、看电视或思考一个深刻的有关存在主义的问题。谎言研究暗示，与多任务处理者对话将导致高语言风格匹配度，常识判断会正好相反。

常识判断再次失败。依拉·陶斯兹克是我实验室的研究生，她有广泛的在线社交媒体学习背景，设计了一个简单的实验来测试这个想法。她让成对的陌生人在单独的实验室计算机上发起的初识对话，作为心理实验的一部分。根据程序设定，随机数字不断出现在计算机屏幕顶部。其中一半的对话中，人们被告知忽略浮现的数字。另外一半的对话中，其中一个伙伴被告知要计算数字“7”出现的次数。然而，和他们聊天的人却被告知忽略出现的数字，并且不知道他们的伙伴在聊天的同时还在计算浮现的数字。换句话说，一半的对话中，两人中的一个是处于分心状态的。

实际上，与谎言研究很相似，与不分心的人相比，分心的人表现出更高的风格匹配度。更奇怪的是，他们往往称更喜欢彼此。在实际词汇使用

方面，分心的学生比不分心的学生更少使用消极词汇和复杂词汇，表达方式更个人化。

在我的职业生涯中，很少有我不相信自己研究结果的时候。和多任务处理者聊天时风格匹配度升高，对我来说是不合理的。所以我把事情交到自己手中，打电话给两个之前的学生，问他们是否愿意参加一个语言项目。项目内容是我们在电话上进行一次非正式谈话，并对其进行记录、转录和分析。谈话实际上是由三段五分钟的片断组成，每段之后，他们需要完成一个简短的问卷。两个人都同意这样的规则。他们不知道的是，三段的某一段中，我会坐在我的办公室里，用最快的速度疯狂地做算术问题。

电话谈话开始，我们聊工作、聊生活、聊我们共同的朋友以及当天的其他问题。在这段时间里，我正在研究算术问题，我记得当时我想："哇，我真的很擅长这样。忙的时候我也能像不忙时那样熟练于社交。"后来，当我转录对话时，我惊讶地发现我在忙于做算术问题时的说话方式有很大不同。我说话结巴且经常重复。面对任何复杂的问题，我都会转移，试图让对方说得更多。类似于依拉的实验参与者，我倾向于更多地笑和使用更积极的语言。

我的两个学生认为，对话的分心阶段和其他部分一样愉快。事实上，LSM值显示，我们在分心阶段的语言使用匹配度等于或高于非分心阶段。然而，最后的结果就是两个学生开始用我和他们说话的方式和我聊天。正如我在心理上远离对话一样，我的对话伙伴也在这样做。令我印象深刻的是，我们没有人意识到这一点。事实上，电话聊天结束后，我和两个学生讨论，询问他们的看法。一个学生说，她模糊地意识到我在做算术问题期间稍微有点分心，但她说："你在电话里经常分心。"噢。

两个互相尊重、互相喜欢的人会学习在对话中如何跟随对方的舞步。他们彼此都将注意力投向对方，即使是在分心或有分歧的阶段。使对话舞蹈变得有活力的是，当一个人心不在焉时，另一个人试图去适应它。然而，在这场舞蹈中一个不言而喻的规则是，两个人最终会专心于此。如果谈话的其中一方或双方根本不关心另一个人，那么对话就有瓦解的危险。

创造一个爱的语言检测器

如果存在一个设备可以警告人们，他们是否合拍不是很好吗？你可以带它一起去约会，晚上结束时，你们两个人会读出爱情检测器上面的“约会合拍指数”，决定你明天是否应该再次相约或者再也不相见。

好消息。我们可能有一个适合你初创公司的原型。确切意义上讲，它不是一个仪器。但是，它只需你用一个录音器转录你的互动，然后通过分析文字评估LSM。噢，为了能让它有效分析，你应该计划以几乎相同的形式约会至少十次。嗯……也许有点像快速约会。

◎约会的步调

事实上，莫莉·艾尔兰、我以及其他几个同事发现，情侣之间在短短四分钟的约会中的LSM可以预测他们是否还想再次见面。正如你可能知道的，在典型的快速约会中，你会遇到8～12个不同的“约会对象”，每个交谈几分钟。每次交谈结束时，每个人在移动至下一个人之前要给另一个人打分。通常第二天，约会者会联系快速约会组织者，告诉他们希望和十个人中的哪一位在一个更放松的场合中相约。大多数人称，有些谈话只是表面话题，有些谈话充满热情，还有一些谈话令人感到疲惫。有时候人们会遇到他们生命的挚爱，大多数时候他们没有遇到。

相亲者可能不总是喜欢快速约会，但研究人员喜欢它。在最近的一个项目中，大约80名访问者允许我们记录他们的四分钟会话，以便我们的研究团队分析他们的语言。简短交流的LSM可以预测这对情侣在未来是否会聚在一起？是的，在一定程度上可以。那些以高于平均水平的LSM为特征的快速约会者，想要未来联系的概率是具有低于平均水平的LSM约会者的两倍。然而，更有趣的是，我们可以预测哪些情侣之后会走在一起，预测效果比个人本身更准。每一个快速约会结束后，参与者会立即完成一个短小的问卷，调查他们对于刚遇到的人的满意度。当然，个人的满意度评价与最后的相遇相关，但是LSM的相关性更高。为什么？两个人最终是否会相聚，取决于双方。可能某个男性觉得一个女性很有吸引力，但是她觉得他很讨厌。两人的步调必须一致。LSM捕捉到的是两人的步调，而问卷只是单独在评估两位舞者。

◎预测年轻人的爱

假设一段关系已经超越四分钟的快速约会，年轻的情侣开始认真约会。

两人间的风格匹配能预测他们关系发展的长期前景吗？初步研究结果表明可以。

大多数热恋中的人都非常关注他们的另一半。他们能觉察到另一个人情绪和行为的微妙转变。语言上，他们的风格匹配度非常高。几个星期或几个月后，一些情侣开始意识到，他们关系的保质期可能有限。随着对另一半关注度的衰减，他们的语言风格匹配度也下降了。

通过了解年轻约会对象的语言风格匹配度，通常可以预测哪些情侣最有可能终成眷属。例如，我和之前的研究生理查德·斯莱彻一起进行了一个年轻约会对象的项目。理查德现在是韦恩州立大学的老师，他想看看情侣通过短信是如何聊天的。他招募了86对情侣，他们说自己每天都会聊天，并同意让我们分析几天的聊天内容。情侣间的短信具有高LSM值和低LSM值，之后我们将这两类情侣进行了对比。

具有高LSM值的情侣占43%，三个月后，其中77%的情侣仍然在一起；相比之下，低LSM值的情侣中，只有52%的情侣还在一起。换句话说，那些自然地将他们的功能词彼此同步的情侣更可能在一段时间后仍保持关系。

通过阅读各种情侣的短信息，我们可以更清楚地看出他们的关系是如何成功或失败的。更成功的高LSM关系表明两个人是如何真正对彼此感兴趣的。另外，短信的整体语气是积极和具有支持性的。具有较低LSM值的情侣更可能会关系失败，通常表现出明显的分离模式。例如，下面这对情侣具有较低LSM值，虽然他们声称其关系充满激情且令人满意。但是，他们的关系失败是显而易见的。

他：嘿，你！怎么样啊？！

她：挺好的，你呢？

他：很好。我和吉姆准备去市里。

她：真好。

他：我得送我妈一幅奶牛的油画作为生日礼物，然后我希望把头发剪了。你在忙什么呢？

她：没什么。

他：我也准备去洗澡了。我们昨晚开了一个温馨的圣诞节派对。我照了照片。很有意思。正在听我的圣诞歌曲，很好听。

她：好的。

他：哈哈，今天话不多啊

她：嗯……确实不多。

他：所以……你周末过得愉快吗？周末做了什么？

她：嗯，还好，工作。

他：啊……好吧，至少你可以赚到钱钱钱！这很好啊。我准备去洗澡了。你想想聊什么!10分钟后回来，如果你还在。吉姆在等我准备好呢。拜拜。

她：那就去吧！

他：想你！哈哈，好吧，拜。

男生话很多，而且我猜，有点儿在假装很开心。他的女朋友很酷，冷漠，而且没什么反应。他的语言个人化，使用高频率的人称代表，而她的语言没有明确意义。通过相对沉默，她在表达自己的烦恼，他避免直接面对。

对比一下，在另一对情侣的对话中，女生正试图在情感上吸引即将成为她前男友的人。前两天会偶尔发短信，女生试图谈论他们的关系，男生不停地忙着作业、锻炼，甚至看电视。在第三天，她终于抓住他可能有时间聊天的机会。

她：你在吗？你能聊5分钟吗？

他：可以，说吧。

她：那你和我聊天的同时在干吗呢？

他：打扫。

她：我很生气，不知道和你说什么。

他：好吧。

她：当你看到那个卡片时，你没有觉得一点难过吗。

他：我当然觉得了。

她：你确定吗。

他：请等一会儿……我确定。

她：还是你只是说说……请快一点儿……还在吗？……怎么回事？

他：好吧……抱歉。

她：你在干吗。

他：我朋友来了。接着说。

她：告诉我是什么让你觉得难过。

他：你以前太好了。但是我还是不觉得我有错。

她：所以你为什么觉得难过……

他：它确实令我难过，但是是你告诉我不要对它有期待。

她：我不明白。我告诉你不要对它有所期待令你感到难过？

他：这样吧，我晚上打电话给你。我没办法这样聊。

第二个例子中，LSM值也偏低。虽然两个人使用人称代词的频率相似，但是计算机分析显示，她的思维比他更具体（使用更多的冠词）、更复杂（例如连词，介词）。而且，她更个人化、情感化，将所有的注意力都放在男朋友身上。而他，却总是分心和回避。

每当她试图靠近或与他交流时，他就会变得遥不可及。这种交互模式在几天的短信聊天中保持一致。然而，当这对情侣回答关于他们关系的问卷时，他们对亲密度和关系满意度的评级都非常高！通常，语言风格匹配度在预测情侣的后期分离方面比他们自己的评判做得要好很多。

LSM方法允许我们通过追踪情侣之间的几次交互来对关系的成功与否做出合理猜测。类似的想法已经在新婚夫妇间进行过测试。华盛顿大学的约翰·戈特曼和他的同事报告说，通过听一对年轻夫妻在实验室中的争吵，可以预测婚姻成功的可能性。戈特曼把夫妻带到他的实验室，让他们讨论有冲突的问题——通常是关于金钱、性或家务的话题。如果在紧张的讨论中，两个成员都表示尊重，试图减少紧张、避免指责，并增添积极的情绪，那么他们的婚姻更有可能持续。另一方面，如果其中一个或两个成员都蔑视对方，主动避免讨论情感话题，或利用任务发起个人攻击，那他们

的婚姻遇到麻烦了。

幸福的婚姻不只是双方使用均等的代词和介词。当两个人以类似的方式和类似的频率使用功能词时，他们在以相似的方式看待他们的世界。但是，有类似的世界观，也不能确保婚姻的幸福。戈特曼的工作提醒我们，稳固的关系也需要共同具有积极情绪这一特征。类似的发现在我们的约会情侣中出现：那些具有高风格匹配度和高比例共享积极情绪的情侣最可能保持在一起。但是有趣的是，他们共享积极情绪不代表长期的幸福或者关系能够一直保持下去。情侣双方必须既态度积极，又彼此关注。

建立一个LSM检测器

想象一下，拥有一个可以追踪所有谈话质量的便携式LSM检测器。也许你可以将LSM检测器指向自己的电子邮件、短信，然后了解你与联络者的连接质量。或者，如果你已经结婚多年，你可以密切关注你们的关系，并在他没有要结束谈话时帮助你解释给你的伴侣。

如前所述，我和我的研究团队已经建立了LSM检测器的粗略版本。登录网站www.SecretLifeOfPronouns.com/synch ，你可以输入你与朋友、恋人或敌人的往来文本。点击按钮，你将收到你们两个人在关于功能词使用方面同步程度的反馈。然后，你可以将你的LSM值与使用该网站的其他人生成的平均LSM值进行比较。更好的是，你可以多次尝试LSM检测器，比较你与不同人的交流。这应该让你感觉到你匹配他人语言的一般技能。

坏消息是，真实世界的LSM检测器具有有限的值。是的，它可以告诉我们何时关系同步，何时不同步，但不能告诉我们同步意味着什么。回想一下，当两个人对彼此充满激情和彼此仇视时，LSM往往会升高。当其中一个人对另一个人撒谎时，一致性同样会升高。所以，如果你的LSM检测器显示结果较高，那么说明你和你的对话伙伴正在密切关注彼此。

我们将LSM检测器视为可以评估会话舞蹈一致性的装置。好的舞蹈需要两个人能够预料到彼此的下一步，某种意义上，能简单地了解其想法。会话舞蹈可能只持续音乐播放的长度，可以反映两个人对彼此感兴趣以及保持想法和行动多年一致的能力。

通过LSM了解过去的关系

使用计算机语言分析的一个优点是，只要某种类型的书面记录存在，我们就可以探索历史上发生过的关系。例如，如果我们可以获得两个人共有的旧信、诗歌或歌词，就可以大概了解一下他们彼此的情感联系。计算机化的文本分析可以阐明我们对历史上存在的婚姻、友谊或联盟的理解。莫莉·艾尔兰已经沉浸在语言风格匹配的研究中，她追踪了几个有启发性的案例，其中包括两个诗人，一个婚姻出了名的幸福美满，另一个则出了名的婚姻不幸。通过分析他们的作品，我们可以更好地洞察每对夫妻的关系。

◎伊丽莎白·巴雷特和罗伯特·勃朗宁

测量语言风格匹配的想法最初用来追踪自然对话中的语言使用。对话应该对思想和共享指示的即时反复是特别敏感的，它们牵涉到镜像神经元的工作方式。通过观察，可以想象语言风格匹配可以发生在更广泛的层面上。如果我听到一个政治对手对一个话题发表演讲，几天或几个星期后，我可能会写出自己对其做出的反应，语言上会匹配原来的演讲。如果我的配偶是一个优雅的小说作家，她的写作风格会影响我写专业期刊文章的方式。可以想象，两个阅读对方作品的专业诗人可能会以类似的方式开始使用功能词——特别是如果两个人相互吸引。一个精彩的例子是19世纪中叶英国诗人伊丽莎白·巴雷特和罗伯特·勃朗宁的情况。

两位诗人第一次通信的时候，她38岁，他32岁，那时他们已经久负盛名。伊丽莎白自14岁以后身体一直很差，她从未离开过家，在写作关于情感、爱情和失去的话题同时，她继续和专制的父亲一起生活。罗伯特是一个充满活力的城市人，他的作品更加偏心理方面，分析性更强。伊丽莎白发表了一首诗，其中引用了罗伯特的一首诗，在这之后，罗伯特写信给她，表达他对她作品的爱慕，更直接地是，表达对她的爱慕，即使这两个人从来没有见过。

接下来的几个月里，他们互相通信，他向她求婚，但最初她有些抗拒，并试图离开他。罗伯特不懈地追求伊丽莎白，最终两个人私奔到意大利，在那里一起度过他们大部分的生活。据说，他们的婚姻极为幸福，两个人都继续发表作品。在他们15年的婚姻中，唯一的黑暗时期是在最后两三年，当时伊丽莎白的身体不行了，导致她在55岁时离世。

在他们的职业生涯中，伊丽莎白和罗伯特都非常多产，产出的诗歌有数百首。即使在遇到之前，两个人使用功能词的方式也往往相似。这可能

解释了为什么这两个人在未通信之前就被对方吸引。从图8-1中可以看出，LSM统计数据表明，在遇到之前和婚姻中的大部分时刻，他们使用功能词的方式是相似的。然而，在一年半的求婚期间，他们写的诗歌发生偏离。在这期间，伊丽莎白在纠结是否要离开她的父亲，纠结她的身体是否会成为罗伯特的负担。

思考这些数据时，你可能很想知道是谁使LSM值随时间变大和减小。总之，LSM被比作两个人之间的舞蹈。我们能否确定哪个舞者对舞步的偶尔蹒跚负更多的责任？事实上，我们可以。仔细观察两个诗人功能词的使用，我们会发现在罗伯特的职业生涯中，他的写作风格令人印象深刻的稳定，而伊丽莎白从一个阶段到下一个阶段倾向于改变使用功能词的方式。例如，他们在一起的大部分时间，伊丽莎白使用功能词的方式倾向于和罗伯特一样。然而，在她被求婚的阶段和生命结束衰退时，她倾向于退缩到自己的世界，写作方式和罗伯特不再那么相似，和她生命其他时候的写作也不那么相似。

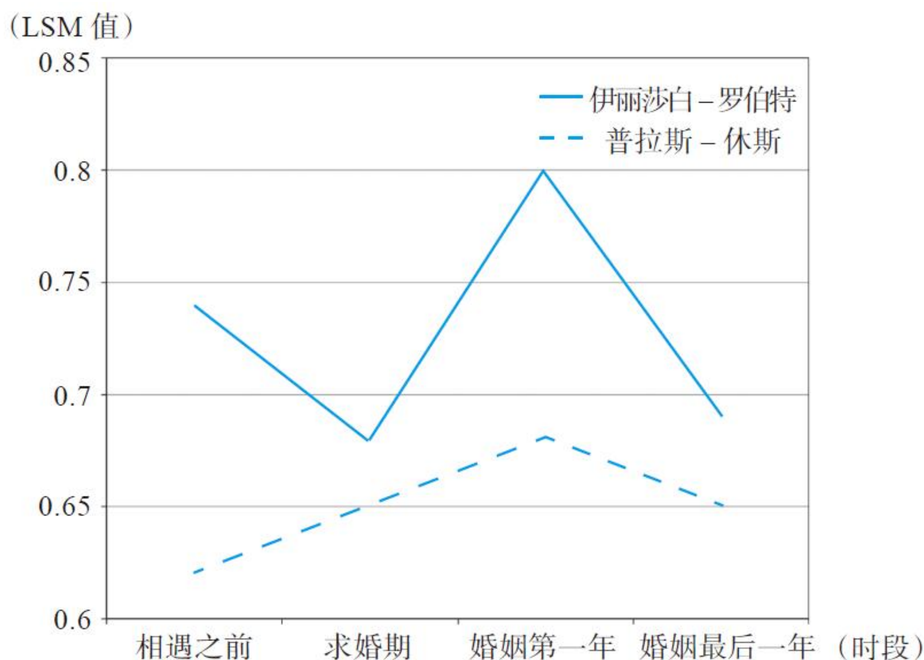


图　8-1

注：伊丽莎白·巴雷特、罗伯特·勃朗宁与西尔维亚·普拉斯、特德·休斯的作品的语言风格匹配（LSM）值。LSM得分越高，两个诗人使用的功能

词越相似。

◎西尔维亚·普拉斯和特德·休斯

勃朗宁的婚姻往往被现代作家所理想化，而西尔维亚·普拉斯和特德·休斯之间的关系则不是。普拉斯出生在波士顿，当时她在剑桥大学获得富布赖特奖学金，遇到了出生在英国的休斯。两人都被认为是诗歌界中新兴的明星。她的作品通常是情感上的、自白的，偶尔异想天开的。他的作品更加理性，通常有关自然和神话。在一个派对上相遇四个月后，他们结婚了：她24岁，他26岁。

婚后的前四年，两人都经历了个人和学术上的成功。相对于普拉斯，休斯的写作生涯很顺利，并赢得了几个奖项。大约五年后，普拉斯的身体状况开始走下坡路，包括几次流感发作、阑尾切除术和一次流产。她很确定地怀疑休斯背叛了她，这经常导致她暴怒。一年内，休斯为了另一个女性离开普拉斯。普拉斯婚前有过抑郁症的历史，这时她陷入了一段沮丧的时期。虽然继续写文章，但是她变得越来越疏远朋友。离开休斯不到一年，在30岁的时候，她自杀了。

比较普拉斯和休斯在他们关系过程中的语言风格很有趣，特别是与勃朗宁夫妇相比。如图8-1所示，普拉斯和休斯在相遇之前有着非常不同的语言风格。她更个人化和直接，相比之下，他更客观、更疏远。在他们相对快乐的岁月里，他们的风格在一定程度上聚合，然后在她生命的最后三年分离。

观察这两对夫妇总体的语言风格有何不同也很有趣。与普拉斯-休斯相比，具有更愉快婚姻的勃朗宁夫妇在他们生活的所有阶段都具有更高的语言风格匹配度。将语言风格匹配总结为勃朗宁夫妇婚姻成功或普拉斯-休斯婚姻失败的原因是一个错误的结论。相反，两个人之间功能词的匹配只反映了两个人往往想法相似。通过观察人们的专业写作，尤其是当两个人可能正在写出针对不同读者的截然不同的话题时，对LSM和人际关系做出大胆陈述是更冒险的事情。

◎爱情之上：弗洛伊德和荣格间的爱慕与蔑视

LSM分析可以告诉我们关于亲密关系的事情。约会情侣间的短信帮助揭示哪些夫妻最有可能成功和失败。更深奥的诗歌分析使我们能够推断已婚诗人思维的相似性，相关地指向他们关系中的高潮和低谷。

但是LSM不只与浪漫的爱和心碎有关。功能词分析可以应用于两个人之间任何种类的关系。只要和其他人有沟通（如信件、电子邮件、微博互

动），我们就可以通过LSM技术来评估相关人员彼此同步的程度。在研究了诗人的语言和婚姻关系之后，我和莫莉·艾尔兰将注意力转向了存在一整套沟通的两人关系。根据你的观点，这可能是爱的关系、认同的关系、未解决的俄狄浦斯情结或者只是朋友关系。

西格蒙德·弗洛伊德和卡尔·荣格之间的关系处于心理学和精神病学的核心地位。19世纪末，弗洛伊德关于精神分析的想法动摇了西方思想的基础。在一系列文章中，他认为人的个性和日常行为是由无意识过程引导的，其中许多与性是高度相关的。他还提出了这样的观点：早期的童年经历塑造了人们未来的心理健康状况。弗洛伊德不仅是一个创造性的思想家，而且也敏锐地意识到他的方法吸引了大众。他担心的一个问题是，他的作品会被认为反映了犹太式思维而被边缘化。

然后是年轻、有事业心的瑞士基督教学者卡尔·荣格。刚从医学院毕业时，荣格被思维障碍的心理学基础迷住，如精神分裂症和无意识的本质。1906年，荣格寄给弗洛伊德他的第一本书，礼尚往来，弗洛伊德寄给荣格他最近的文章。

通了几封信后，他们成为亲密的朋友。大多数学者认为他们非常喜欢对方，但同时他们也看到了密切关系的专业优势。事实上，在他们第一次见面后，弗洛伊德在一封信中表明了荣格是他“亲爱的朋友和继承人”。到1908年，荣格能够写信给弗洛伊德说：“让我享受你的友谊，不是平等的关系，而是像父亲和儿子那样。这样对我来说是合适和自然的。”

1906～1913年，两个男性之间至少通信337封。这些年里，弗洛伊德的声誉飞涨。荣格形成自己的力量，1911年，两个人之间的紧张关系开始建立。荣格觉得弗洛伊德过于强调性行为；弗洛伊德感到荣格不忠于他的观点。在通信的最后几个月里，荣格指责弗洛伊德傲慢和头脑封闭。弗洛伊德回应道，他们应该“完全放弃……个人的关系。”

通过分析弗洛伊德和荣格信中的功能词，我们发现了他们语言风格匹配方面的可预测模式。正如你在图8-2中看到的，他们的风格匹配度在前四年特别高，但之后骤降。更密切的分析发现，两个人在第一年对这段关系的投入程度一样。然而到最后，弗洛伊德使自己脱离，他的语言变化更多。事实上，在他给荣格的最后一封信中，他建议他们不再做朋友，他还说：“我不会因此失去任何东西，因为我和你之间唯一的情感纽带一直是一条细线——这是过去的失望挥之不去的结果。”

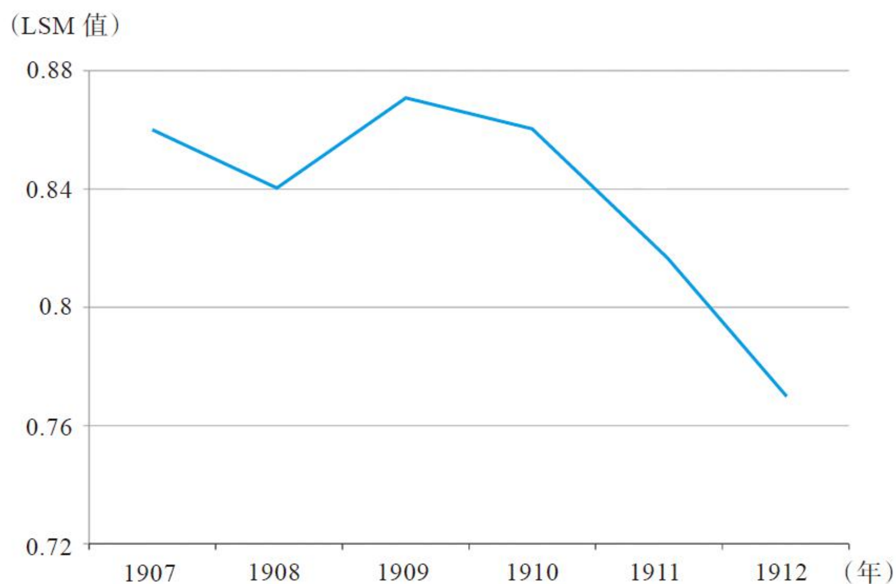


图　8-2

注：西格蒙德·弗洛伊德和卡尔·荣格信中的语言风格匹配度。值越大，两个人在他们的信中使用的功能词就越匹配。

通过LSM了解亲密关系

会话舞蹈在许多方面发挥作用。谁曾想到，人们使用代词、介词和其他功能词的方式可以告诉我们这么多关于他们关系的事情？

用一点儿计算机魔法，我们可以使用LSM作为显示两人社会联系的计量器。或者更具体地，作为人们彼此同步程度的标志。

说两个人同步到底是什么意思？从最基本的层面来说，人们同步功能词的使用说明他们互相关注。他们可能不喜欢彼此，他们可能不信任彼此，但他们正在观察和倾听彼此。幸运的是，人们避免与他们的敌人花费太多时间，而选择与他们喜欢的人交谈。大多数好朋友或恋人间的对话的特点是高LSM，原因相同：他们正在关注对方。

很难理解两个人如何能够快速适应彼此语言风格。正如我们所看到的，人们通常在与陌生人交谈时只需几秒钟时间适应。

两个人立即适应彼此的正式性、具体性、情绪性程度和思维方式。两个人都在不断追踪哪些代词指向哪些人，以便在脑海中保持记录，记住“她”“他”或“它”指的是什么。为了保持对话持续不断，两个人必须适应转变的主题。事实上，如果其中一个对话伙伴暂时分心或开始表现异常（说谎时），另一个人必须投入更多的能量来维持互动。

我们很容易理解为什么LSM在爱情对话和激烈的争吵中都如此之高。即使在互不关心的人之间的最无聊的讨论中风格匹配也惊人的高。幸运的是，镜像神经元系统不断监控说话者所说的内容，并帮助说话者以最小的努力传达需要说的内容。当然，有时候，风格匹配彻底失败。例如，有些人很难做到语言风格匹配。某些疾病（如自闭症）会干扰人们模仿的能力以及与他人进行社交联系的能力。一些研究人员认为，因为镜像神经元系统功能中断，所以自闭症谱系障碍，包括像阿斯伯格综合征这样的诊断会阻碍人际互动。

在现实中，我们不是总能够理解我们的对话伙伴。多数情况下，一方或双方根本对谈话或正在谈话的人不感兴趣。有时，一个对话伙伴根本不想听到对方要说的话。例如，我最近在一个会议上，坐在两个人旁边，让我们暂且称他们为大卫和艾哈迈德。在会议之前，大卫对艾哈迈德的妻子做了一些毁谤性的评价，而艾哈迈德了解到了这一点。大卫不知道艾哈迈德知道了这件事，开始友好地聊他们的一个共同好友。在那段时间内，艾哈迈德避免与大卫目光接触，大卫的故事一讲完，艾哈迈德会接过会话的主动权，开始与我谈论最近看的一本书。每当大卫试图加入对话时，艾哈

迈德都会忽略他，或者再次改变话题。后来，当我问大卫对话情况时，他完全没意识到艾哈迈德忽略了他说的一切，并且认为对话很顺利。

艾哈迈德和大卫的对话提醒我们，人们并不总是能够意识到他们未能与他人取得沟通。很像约会情侣的研究，LSM值测量显示，人们可能完全没有理解对方的意思，却认为他们的互动是顺利的。我们的研究发现，这些不正常互动中的伙伴往往不能理解问题的所在。也许，另一个人巧妙地拒绝他们这件事太过危险而难以承认。也许忽略另一个人的人也看不到这个过程。除非两个人都随身携带LSM检测器。

第9章　通过语言了解团体、公司和社区

越多团队成员说话相似，团队的凝聚力越强。如果一个团队中的人以相似比例使用相似的功能词，则团队表现更好。

管理顾问有时会区分我-公司，我们-公司和他们-公司。为大致了解公司的氛围，他们让员工聊一聊他们典型的工作日。如果员工提到公司时说“我的办公室”或“我的公司”，那么说明工作场所的气氛通常很好。在这些我-公司工作的人是相当快乐的，但不是特别注重公司本身。但是，如果他们提到公司时说“我们的办公室”或“我们的公司”，值得关注一下。我们-公司里的人把工作场所当作自己的一部分。这种感觉也许解释了为什么他们工作更努力，员工流动率更低，以及他们对工作有更大的满足感。但是请注意，如果一个公司的员工开始称之为“公司”，或更糟，称“那家公司”，并称他们的同事为“他们”，他们-公司可能是噩梦，因为员工声明他们的工作身份与他们无关。难怪咨询者报告说，他们-公司有不愉快的员工和高流动率。

人们如何谈论他们的公司只是探索一个组织的团体动力的一种方法。通过听人们在某个团体中使用的语言，可以揭示关于团体内部工作的几个特征。通过电子邮件、网页、会议记录和其他文字线索，我们可以衡量团体成员想法类似的程度，也可以衡量他们的凝聚力、生产力、正式度、共同的经历，以及在某些情况下，可以衡量他们的诚实度和做出改变的意向。

本章可能与一些人不相关。如果你没有任何亲戚、同事或朋友，或不认识任何一个单位、社区里的人，你可以跳过这一章，直接读下一章。其他人都应该继续阅读。

“我们”作为身份标识

几年前，我在飞机上遇到一个咨询员，从咨询员那里我了解到我-公司、我们-公司和他们-公司。大约在同一时间，情侣研究者发现了一些类似的模式。在一个典型的研究中，已婚夫妇被邀请进入实验室，并鼓励谈论他们的婚姻或关系中的一个问题。有时会有火花飞溅——有时是爱的火花，有时不是。

一般来说，一对夫妻在被访问关于他们婚姻的事情时，使用“我们”越多，越好。当夫妻使用温暖的、模糊的“我们”，即指“我的配偶和我”时，它表明健康的关系，在一些研究中，它能预测他们的婚姻将持续多久。有趣的是，使用“我们”会预测良好的关系，只有夫妻双方在访问者面前才会这样。其他研究要求已婚夫妻几天内佩戴数字录音机，研究没有找到任何“我们”模式。在另外一个人面前使用“我们”通常意味着“你”或“只有你”。一对夫妻在与第三方谈话时使用“我们”表明关系令人满意。然而，夫妻彼此交流时使用“我们”，则很少能预测他们关系的质量。

实验室里，当谈到婚姻分歧时，“我们”表明良好关系，而使用“你”表明出了问题。使用“你”的话，例如“你”“你的” 和“你自己” ，在有毒的谈话中最为明显——通常是两个参与者互相指责对方的各种缺点。

“我们”甚至可以拯救你的生命。在一个项目中，患有心力衰竭的患者与他们的配偶一起被访问。他们被问到一系列问题，包括“当你想起你们两个人如何应对心脏病时，你认为你们做到最好了吗？”

配偶在他们的答案中使用“我们”越多，六个月后，病人的健康状况越好。配偶使用“我们”表明他们认为伴侣的健康问题是共同的问题，两个人都致力于解决它。当两个人一起努力应对疾病时，会减少患者的身体和情绪压力。

即使你完全健康，“我们”也许也会拯救你的生命。对商业航空公司驾驶舱记录的分析发现，飞行机组之间缺乏沟通是20世纪所有航空公司一半以上事故的促成原因之一。某些情况下，飞行员建立了一个阻止异议的不利氛围。其他情况下，一个或多个机组成员分心，无法听取他人的关键信息。一个反复出现的主题是，最有效的机组人员是那些紧密相连，并且觉得他们是一个团队人。例如，布莱恩·塞克斯顿和罗伯特·海姆里奇分析了扩展模拟飞行期间机组的语言。机组使用“我们”越多，错误越少。在分析航空器碰撞的驾驶舱记录时，与由不可避免的机械故障引起的事故相比，具有明显人为差错的事故与较低频率使用“我们”相关联。

“我们” 和“我们的” 这样的词可以成为强大的身份标识。当人们向陌生人讲述“我们的婚姻”“我们的企业”或“我们的社区”时，他们正在公开声明他们是谁、他们与谁相关。例如，“我们的婚姻”是一个共享和联合体。同样，“我们的业务”和“我们的社区”是构成我们的一部分。

扩展的“我们”：人和团体

人们对自己的看法是不断变化的。“我” 和“我们” 之间的转换可以非常微妙，几乎可以瞬间发生。在对话中，说话者和听话者都不会意识到听到了哪些代词。然而，说话者使用“我们”来指听众和他们自己具有心理和社会意义。

下面是偶然听到的两个人在谈论房地产交易时的对话，你可以感觉到这种转变。亚历克斯大约45岁，是一个涉足房地产交易的律师。大约40岁的利兹·安成功地在股市顶峰时期脱离，并参与了一些项目的投资。

亚历克斯： 我有一个交易，你可能会感兴趣。它涉及在橡树街购买那套房产。

利兹·安： 我现在最不想做的就是和达顿集团进行另一场战争。

亚历克斯： 不是达顿集团。原始持有人出于税务目的要买它而联系的我。

利兹·安： 他们要价多少？我的风险是什么？

亚历克斯： 算上私下交易，可能350。我们可以报价300。

利兹·安： 我不确定这是一个合适的时机。但是如果我们把达顿转到次要地位，我会很开心。

显然，两个人相互认识，而且一起工作过。简短对话开始时，两者是分开的个体，有着自己的计划。他们说话时，每次都用“我”字。接近最后一行的时候，亚历克斯提出“我们可以报价300”。亚历克斯眼中，两个人因为共同目标从两个人微妙地变成一个团体。利兹·安最后一句话表明，她认同他们的共享身份。虽然亚历克斯可能没有意识到利兹·安有使用“我们”，但是他的大脑可能检测到她倾向于投资橡树街的资产。

使用“我们”通常表示一个人感觉自己是一个团体的一部分。有经验的销售工作者经常关注人们在使用“我们”方面的转变。正如房地产对话表明的那样，当客户开始用“我们”来指代销售人员与客户的关系时，说明产生了一种重要的情感纽带。如果你是销售人员，你能通过自己使用“我们”来加快这种关系吗？估计不能。过早使用“我们”的话，很像政治家的语言，通常被认为是不真诚且操纵人的。

“团体感”往往是虚幻的。有时候，人们感觉自己是所在团体中的坚实一部分，而在其他时候，虽然围绕着完全相同的人，他们可能感到孤独、疏远或寂寞。通过追踪人们对“我们”和“我”的使用，可以检测到他们对团体身份的看法。相同的语言分析也可以告诉我们团体本身的东西。

◎与日俱增的团体感：从“我”到“我们”

与他人交谈时间越久，人们使用“我们”会越多，使用“我”越少。当我们了解别人的时候，会放松警惕，开始接受他们。“我们”增加和“我”减少的模式出现在广泛的团体中。

快递约会是一个完美的起点。回想一下上一章中的快递约会项目，陌生人连续约会十个人，每个四分钟。即使在这些荒唐的短暂见面中，人们在“我们”和“我”使用上的转变也很快，而且可预见。在分钟计时的短暂的快递约会中，两个人“我”字的使用都减少了，“我们”的使用增加了。

“我”-减少/“我们”-增加效应可以在约会环境之外看到。在一些研究中，学生来到心理学实验室，在一个“开始了解你”的背景下与一个完全陌生的人聊了15分钟。

前5分钟，两位参与者通常都会谈一些他们的背景、专业，以及他们住在哪里。因为两个人都在谈论自己，所以往往使用“我”的频率会很高。在接下来的5分钟，他们会减少谈论自己，因为他们开始建立彼此之间的共同点。最后5分钟，和最初相比，他们使用“我”字的频率下降了10%～50%。同时，他们使用“我们”的频率增加了20%～200%。这种模式在两个陌生人的在线聊天中甚至更明显。

同样的转变发生在更大的团体中。商学院实验中，四人为一组，一起在30分钟内根据任务做出团队决策。出现了同样的结果：前10分钟到最后10分钟，团体中人们使用“我”的频率降低了19%，“我们”增加了39%。

更有趣的是，现实世界的团体中，“我们”和“我”的转变会持续几个小时、几天、几个月，甚至几年。前面讨论的航空公司驾驶舱项目发现了类似的模式。机组人员在一起的时间越长，使用“我们”越多。

“我”-减少/“我们”-增加效应可以在更大的时间范围内扩展到更大的团体中。一个项目涉及十几名工程师、经济学家和计算机专家，他们从事一个复杂的在线国防项目有近两年的时间。另一个项目涉及大约20名专业治疗师，作为专业培训的一部分，他们三年内每年两次会面。另一个项目追踪披头士乐队在一起十年的歌词。对于所有这些团体，“我”的使用在逐月下降，而“我们”的使用在增加。

所有这些意味着什么？我们与其他人一起度过的时间越多，我们的身份与他们融合得越多。我们可能不一定喜欢或信任他们，但是随着经历交织在一起，我们把自己看作同一个团体的一部分。随着我们年龄的增加，这种现象有一个有趣的推论。总的来说，年龄越大，我们与周围几乎每个人一起度过的时间越多。

人们甚至可以预测，老年人会比他们年轻时使用更少的“我”和更多的“我们”。这是真的。在几乎3000人的语言研究中，70岁以上的人使用“我们”的频率比青少年高54%，“我”的频率比青少年少79%。

◎把“我们”聚在一起：事件创造团体认同

逐渐接受团体身份是一个我们很少注意到的自然过程。谈话中，前一分钟，我们的脑海里还是“我”，后一分钟就变成了“我们”。它就这么发生了。然而，还有其他时候，我们的团体认同方式会有快速和戏剧性的改变。

最简单的例子是我们对运动队的忠诚。最富智慧的社会心理学实验之一，是由罗伯特·西奥迪尼和他的同事于20世纪70年代中期开展的实验。足球赛季期间，所在大学拥有排名靠前足球队的学生接到电话调查，据称是处理校园问题。在前几周，他们本校的足球队赢得了一场大型比赛，但也输了一场大型比赛。几个基本问题之后，采访者问了赛季中两个关键赛事其中一个：“你能告诉我这个比赛的结果吗？”

如果本校的球队赢了，他们通常回答：“我们赢了。”但如果本校的球队没能赢得比赛，他们的答案更有可能是：“他们输了。”因球队的胜利而得到部分荣誉是沐浴在他人荣耀中的一种形式。基本上，我们想要靠近成功的团体，而远离失败者。

希望属于一个强大团体的意识可能根源于进化。大多数社会动物在受到威胁时会寻求群体的保护。甚至外部人的出现，都能使人们更加意识到自己的社交网络。同一个足球项目中，西奥迪尼的采访者有时声称他们和受访者住在同一个城镇。其他时候，采访者声称他们来自国外。换句话说，一半时间，参与者认为他们正在和自己相似的人谈话，而另一半时间，他们相信自己正在接受一个外国人的足球采访。当外人交谈时，“我们”的效果更强烈。当想起团队竞争时，我们是成功团体的成员身份变得更加重要。

当一个外部团体威胁到我们自己团体的存在时，成为团体一部分的需要最有可能发生。没有什么能比对生命的攻击更快地刺激人们。美国“9·11”事件提供了有力的案例。9月的第一天，乔治W.布什当上总统不

到8个月，支持率约为54%。在“9·11”事件发生后的一个星期内，他的支持率暴涨至近90%。在国家的每个城市，人们挂起了美国国旗，所有的新闻机构都在讨论流露出的爱国主义和民族自豪感。

你可能还记得我们对“9·11”事件之前两个月到事件结束后两个月的上千条博文进行的大规模项目。在被攻击的几分钟内，LiveJournal.com网站的博客使用者开始从“我”转换到使用“我们”。事实上，这种模式持续了几个星期。“我们”上升的同时伴随着积极情绪词汇的短暂下降，但随后积极情绪词汇长时间处于增多状态。类似“9·11”的可怕创伤会产生令人意想不到的效果，它把人们聚集在一起，使他们更少关注自我，几天后，使他们更开心。

“自然和人为的灾难可以使人们聚集在一起”不是一个新想法。当伦敦居民在第二次世界大战中长时间遭受德国的夜间轰炸时，显示出人民之间更加紧密的社会纽带。一个经常被引用的统计数字是，这段时间内自杀率急剧下降。最近的数据显示，2001年9月11日后几周内，以及在2006年地铁爆炸后的英国，自杀率大幅下降。你会记得之前讨论过类似的发现，有关1980年圣海伦火山喷发事件。那些损失最严重的城市居民后来称，他们很高兴生命中发生这样的事，很高兴火山喷发把社区人聚集在一起。

更多的研究发现了相同的效果模式，其中许多已经在本书其他地方提到。“我们”-上升和“我”-下降现象如同税收和死亡一样可靠。示例如下。

·报纸文章、给新闻编辑的信，以及在1999得克萨斯州农工大学造成12名学生死亡的篝火悲剧的网上悼念词。灾难激发学生产生比以往更紧密的社会关系。与该地区所有其他大型大学不同的是，该学校之后6个月的疾病率与前一年相比下降了40%。

·在线聊天组谈论1997年戴安娜王妃在车祸中死亡的事情。模式持续了大约一周。

·在线论坛讨论1995年168人遇难的俄克拉何马城爆炸事件。即使在早期共享在线交流的日子里，人们也总是急于谈论自己的感受。

·卡特里娜飓风摧毁城市四个月后，新奥尔良市大学生的数百篇文章。

·2004年3月11日的马德里地铁爆炸案发生后的几个星期，对西班牙人的数百个访谈和调查。

·伦敦地下爆炸事件之后的英国，2007年贝娜齐尔·布托被暗杀后的巴基斯坦，2008年孟买发生袭击后的印度，以及2010年总统和其他领导人

遇难于空难后的波兰，谷歌网络搜索均显示出类似的模式。

使用“我们”和团体凝聚的感觉不是在人们受到威胁时才发生的。就像足球的例子，当团体成功或令人钦佩时，人们接受他们的团体身份。谷歌实验室开发了一款名为“谷歌趋势”的精彩应用程序，允许用户在搜索时追踪其他用户使用的字词。有了它，任何人都可以看到特定城市、州或国家随着时间推移是如何使用词汇的。使用“谷歌趋势”可以很容易地看出在文化自豪时期“我们”使用增加的例子。全国范围内，2008年巴拉克·奥巴马获选后，以及2010年英国全国选举中保守党取得成功后，“我们”的使用增加。在加拿大，2010年2月冬季奥运会期间，“我们”的使用达到高峰。

◎团体认同的重要性和局限性

“我们”的使用显示出团体认同。通过仔细分析它们，可以了解说话者与其他人的联系。知道某人是否认为他是婚姻、家庭、公司、社区或国家的一个组成部分，这是宝贵的信息。事实上，许多心理学家和社会学家致力于理解团体认同。一个人认同团体，只是更加忠诚于这个团体。他们倾向于依靠团体中的人，而不信任团队之外的人。团体身份常常与模式化形象、偏见和歧视有关。许多已经发生的家庭仇恨、地区战斗、种族灭绝事件和世界大战中，团体身份起到了战斗口号的作用。

如果你对团体和它们的运作方式感兴趣，那么了解团体成员的身份是一个好的开始。但这只是故事的一部分。

我是各种类型团体的成员并不意味着我必须喜欢每个团体中的所有人，能与他们相处得很好，或甚至想和他们一起吃午饭。真正理解团体如何运作，追踪他们如何相互沟通非常重要。

团体认同之外：捕捉团体动态

语言演变是为了让我们可以与其他人沟通，这点很容易被忘记。要了解一个团体如何运作，看看团体中的成员如何相互作用。当团体中有人说话时，团体的其他成员是否在听？当他们在一起处理问题时，常常以相似的方式说话吗？

在有关密切关系的章节中，我们讨论了LSM。LSM通过查看两个人使用隐性词（人称代词、冠词、助动词等）的类似程度来衡量他们彼此是否同步。一对情侣在谈话、电子邮件或短信中使用隐性词汇的方式越相似，这对情侣越“合拍”，几个月后更有可能仍然在一起。和情侣一样，所有团体在凝聚、同步或合拍的程度上都会有所不同。你可能最近和一群朋友一起吃晚饭，并对团体的完美和谐感到惊叹。一个星期后再次见面，同一群人可能没有明显的原因而感觉完全不同。我们大多数人在工作上、家里或在线聊天中都有类似的起伏。通过测量每个人在这些互动中使用隐性词汇的相对情况，我们可以开始了解团体的运作情况，甚至可以确定问题可能出在哪里。

◎小型工作团体：团结和生产力

想象一下，你被扔进一个工作团队，必须提出一个复杂任务的解决方案。作为商学院练习的一部分，你的团队必须决定，购买LuProds制造公司是否是一个好的投资。你了解到，LuProds设计了一个新的自行车齿轮系统，可以改革自行车行业。在你的五人工作团队中，没有人了解LuProds或自行车行业。然而，在接下来的四个小时内，你的团队必须撰写一份报告，其中列出有事实依据支持的建议。为了增加压力，还有其他十个团队在执行相同的任务，并且必须确保你的团队是最好的之一。

要在这样的任务中成功，每个成员必须承担一个独立的任务。必须有人了解自行车业务、齿轮业务、LuProds的经济历史、竞争对手、其管理团队的质量，以及潜在的其他LuProds购买者。团队成员必须以协调的方式密切合作，根据所有可用信息提出一个总体建议。

一个好的工作团队具备什么？即使是你的团队在一起的最初几分钟，我们也可以开始预测它将如何做任务。这可能不会让你感到惊讶，但是团队成员彼此使用功能词的方式是最好的预测器之一。这个项目的变体已经在世界各地的商学院和心理学系所进行。在两项研究中，我和同事计算了团队中每个人的LSM值。因此，例如，如果团队中有五个人，我们可以确定五个人中的每个人与该团队平均值的匹配程度。每个团队的平均风格匹

配值或LSM值初步衡量了整个团队以相似方式说话的程度。越多团队成员说话相似，团队的凝聚力越强。换句话说，高LSM值反映了团队的强凝聚力。

更重要的是，风格匹配与团队的表现有关。如果一个团队中的人以相似比例使用相似的功能词，则团队表现更好。即使在团队提交答案前，我们的语言分析也可以对不同团队的最终成功做出合理的预测。至少在受控的实验室环境中，团队成员使用的功能词（例如代词和介词）表明，参与者达成了共识，即他们对任务和需要做什么有共同的设想。

◎大型现实团体：维基百科

团体行为的实验研究告诉我们，人们在更自然的环境中如何作为。你的生活中有多少次会与由四个20岁的人组成的团队一起做一个需要四小时完成的、你毫无了解的商务主题的任务？语言风格匹配在真实的现实世界中能够预测成功吗？答案在维基百科中。

维基百科创建于2001年，是一个类似在线百科全书的信息来源，有几百万篇文章。许多文章是精心编写的，已经由几十个，有时由几百个人仔细编辑过。对于最常被阅读的文章，这些人进行了详细的非正式复审。通常，一个人开始有关特定主题的文章。如果对话题感兴趣，其他人将访问该网页，并时常更改原始文章。在许多方面，维基百科的文章实际上是两篇文章。随意浏览的游客只看到最终成品。然而，通过点击“讨论”选项卡，可以找到各个参与者之间的对话。这些互动往往是相当专业，有时详细和透彻，但偶尔也会坦率到粗鲁和令人讨厌。

令人惊叹的是，维基百科反映了一种现实运作的知识民主形式。虽然很多学者不愿承认，但维基百科经常是他们了解一个新话题时访问的第一个专业网站。如果一个人对团队运作感兴趣的话，它也是一个丰富的语言数据来源。事实上，它吸引了我的研究生依拉·陶斯兹克，她获得了得克萨斯大学奥斯汀分校超级计算机团队的帮助，下载了所有的维基百科，包括后台的对话，我必须说，这是一个非凡的任务。

开始时，依拉确定了大约100个拥有相对广泛维基条目的美国城市。这些是中等城市，人口从50万到150万人不等。多年来，每个城市的网页已经由至少50个不同的人编辑过多次，所以关于条目的讨论很多。包括文章本身在内，每个条目都经过了复审过程。事实上，所有的文章都被维基百科归类，从“残端”（意思是不值得称为条目）变成典范。

各个编辑在彼此交流中使用相似语言的程度反映了更高质量的文章。正如在实验室研究中，维基百科彼此同步的作者和编辑团队通过功能词使

用衡量写出了最好的、最具权威的文章。现实世界中的语言风格匹配反映了更好的现实世界的作品。

◎更大型现实团体：城市和克雷格列表

如果我们可以查看一个不断变化的维基百科编辑团队的凝聚力和相互联系，为什么不观察一下真实的社区？一些城镇的内部联系比其他城镇更紧密。我们大多数人偶尔会来到某个城镇或社区，在这里，人们似乎在观念、食物偏好和社会行为方面非常相似。而在其他地方，几乎没有人认识他们的邻居或对他们社区的历史或价值观有感觉。

维基百科项目显示，可能通过简单地测量社区中的人们如何使用语言，评估人们之间的联系。另一个惊人的互联网发展是网站Craigslist.org（克雷格列表）的成长。在美国大多数城市里，克雷格列表允许人们购买、出售和送出的商品和服务的数量令人目眩。

广告本身是免费的，人们可以写得尽可能多，甚至可以插入图片。许多美国印刷报纸的消亡部分归功于克雷格列表，因为人们更可能在克雷格列表上出售他们的汽车、租赁他们的房子或赠送他们的小狗，而不是使用报纸上付费分类的广告。2008年，我和我的学生保存了美国30个不同的中型城市里所有克雷格列表的广告。因为广告类别过多，所以我们只看了汽车、家具和室友征集的广告。每个类别中，我们分析的广告数量在6000到10000之间，并评估了不同城市的人们是如何使用功能词的。更具体地说，我们很好奇不同城市使用代词、介词、冠词等功能词的方式有多相似。

结果是令人着迷的。在几个城市中，人们倾向于使用与邻居相同的语言风格来撰写自己的广告。在波特兰，人们通常以一种具有轻微负面情绪的个人风格写作。盐湖城的人们在广告中则格外乐观。虽然在波特兰和盐湖城的克雷格列表作者可能有独特的风格，但这两个城市中的人在广告中以同样的方式思考。相比之下，加利福尼亚州的贝克斯菲尔德和北卡罗来纳州的格林斯伯勒的居民在他们的写作风格上更加无规律——意思是，有些人的写作可能非常正式，其他人则非常不正式。他们的风格根本不一致。

使用类似于语言风格匹配的方法，我们能够计算任何给定城市中的居民使用语言的相似程度。研究结果表明，城市居民使用的语言越相似，城市的凝聚力越强。

当你看到表9-1列出的十个语言上凝聚力最强和最弱的城市时，它们之间的差异可能不会立即显现。这不是北方与南方、自由主义与保守主

义、富裕与贫穷的关系，也无关城市的种族构成、年龄分布或城市移民情况。相反，两个名单之间最显著的区别之一是，最强凝聚力城市比最低凝聚力城市拥有更高和更平等的收入分配。使用被称为基尼系数的统计数值，人口统计学家可以确定财富在任何给定的城市、州或国家的传播程度。居民以类似的方式使用语言的城市往往是在收入方面更相似的城市。城市的贫富差距越大，克雷格列表广告的写作风格就越不一样。

表　9-1

语言上凝聚力最强城市（前十名）	语言上凝聚力最弱城市（后十名）
波特兰，俄勒冈州	贝克斯菲尔德，加利福尼亚州
盐湖城，犹他州	格林斯伯勒，北卡罗来纳州
罗利，北卡罗来纳州	路易斯维尔，肯塔基州
伯明翰，亚拉巴马州	俄克拉何马城，俄克拉何马州
罗切斯特，纽约州	代顿，俄亥俄州
哈特福特，康涅狄格州	埃尔帕索，得克萨斯州
新奥尔良，路易斯安那州	杰克逊维尔，佛罗里达州
里士满，弗吉尼亚州	哥伦比亚，南卡罗来纳州
伍斯特，马萨诸塞州	塔尔萨，俄克拉何马州
图森，亚利桑那州	奥尔巴尼，纽约州

注：凝聚力通过各个城市居民在克雷格列表广告中使用功能词水平的可比程度计算而来。

你可以看到为什么基尼统计数值是重要的。

在富人和穷人差距更大的城市中，城市的不同元素不太可能相互交流。随着一个城市的收入差距变小，居民可能会有更多的共同点，并与他人交流更多。并非巧合的是，他们还应该有更多的物品和服务想在克雷格列表上和同城的人交换。

这个项目的魔力不在于城市的收入分配和社会模式之间的联系。相反，它显示了语言如何在最平凡的地方揭示关于城市社会关系的重要信息。不管是家庭、工作团队、公司还是整个城市，所有团体的成员在彼此交流时，语言中都会留下社交活动和心理活动的痕迹。语言是人类创造的

元素之一，它将我们的思绪和想法连接在一起。通过追踪语言，我们可以对社交网络产生一个概念。

团体定位：语言如何显示一个团体在做什么以及它在哪里

考虑一下目前为止我们了解到的内容。个人使用词汇，特别是代词，来表示他们对各种不断变化的团体的归属感。“我们”的测量反映了团体身份感，而一致性测量，例如语言风格匹配度，反映了团体成员共同享有的世界观。

综合起来，各种发现指向大多数团体的共同心态。无论是家庭、学校、工作团队，甚至社区，人们在任何团体内都会很快开始以相似的方式谈话。在小的团体中，他们彼此讨论相同的主题，且某种意义上，他们使用的功能词相似不是太令人震惊的事情。但是随着团体规模的扩大，并不是每个人都一定会与其他人交谈，令人印象深刻的是，团体仍然保持自己的语言风格。

社会语言学家不会感到惊讶的是，大型的团体有着相似的语言风格。现在有一系列重要的研究，研究的是一个社区、城市和整个地区是如何发展自己独特口音的。此外，人们通常可以根据谈话对象来调节口音。如果我的妻子在家里打电话，我通常可以通过她的节奏、音量、音调，甚至口音来猜测她在和谁通电话。如果我的大脑回路有所不同，我也许能听出她的功能词的使用方式，并做出更好的猜测。

想想功能词使用“传染性”的影响。如果功能词真的和口音有点像，那我们的语言应该取决于自己所在的团体，取决于团体可能做的事情，甚至可能取决于团体所在的地方。我们可能意识不到，但我们可能会根据恰巧所在的团体和团体正在做的事情而改变自己使用的功能词。

总的来说，每个团体都有自己的语言指纹。语言指纹的一部分可以揭示该团体一起工作的良好程度及其成员与该团体认同的程度。语言指纹的其他部分反映了团队的想法和感觉，以及情感上开放、正式、疏远的程度等。理论上讲，任何团体的语言都应该为团体及其成员的动态提供线索。

◎公共场合捕捉语言：语言的社会生态学

当然，团体有不同的目的。有时候，我们会在一个完成任务的团队中、游戏团体中，或者只是在社交。不同的任务需要不同的说话方式。通过分析不同团体需要的语言，我们可以获得对团体本身性质的新看法。

几年前，我和我的学生携带数字录音机出入各种公共场所。我们会高

举着我们的录音机慢慢穿过餐厅、拥挤的走廊、邮局、大学篮球场和幼儿足球比赛现场、教室、游泳池。“词捕捉”项目几乎没有规则，除了我们不得不一直移动，不能在任何给定的会话中停留超过几秒钟。目标只是在不同类型的场合捕捉词汇。多年来，我们收集了更多的语言样本，包括团体中的人和第一次见面的陌生人。每个语言样本都被转录并添加到我们的“词捕捉”进行存档。

“为什么要做这样一个棘手的项目？”我的很多朋友和同事经常问。我想知道人们是否以及如何根据他们所处的情况改变使用语言的方式。结果很清楚：每个语境都有自己的语言指纹。

花一分钟，思考“情境语言”表中的发现（见表9-2）。每个一般情况都与一种非常不同的语言模式相关联。人们在吃午饭或晚饭坐着和他人交谈时，或者是坐在公园里与他人交谈时，他们使用“我”字的频率较高，与此同时，第三人称、过去时态和否定词用得很少。这听起来可能有点儿难懂，但实际不是。这些语言模式表明了叙事性。而且事实上，在阅读转录文字时，我们可以清楚地看出，坐着和他人交谈的人都在讲述自己和他人的故事。

表 9-2

情境语言			
情境	高频率使用词汇	低频率使用词汇	诠释
坐着吃饭，咖啡厅，公园长椅上	第一人称单数、第三人称单数、过去时态、否定词	冠词	叙事、缺少细节
体育赛事，打球	第二人称复数、短词、非人称代词、现在时态	第一人称单数、将来时态、连词	当下、和团体一起、简单的生活
工作团体	第一人称复数、非人称代词、大词、冠词、介词、连词、消极情绪词汇	第一人称单数、第三人称单数、过去时态、副词	复杂的、分析的、严肃的、和团体一起
移动中的人们：走廊里、邮局	第一人称单数、第二人称单数、动词、现在时态、积极情绪词汇	冠词、介词	简短、表面、当下、愉悦
和陌生人聊天	副词、限定词、积极情绪词汇	第一人称复数、非人称代词、过去时态、否定词	积极的、描述的，但是没有共同的过去

在运动场合，语言更加乐观，缺乏内省。无论是打台球还是观看体育赛事，男性和女性都沉浸在比赛中，而不是专注于自己。他们那一刻很享受，仿佛是团队的一部分。这可能就是运动的魅力——逃离自我。工作团体的语言有一个有趣的对应。他们也使用高频率的“我们”和低频率的“我”。最大的区别是工作团使用更复杂的语言，并倾向于表达更多的负面情绪。

人们在走动时与他人交谈往往是个人的、乐观的和表面的。“嘿，你怎么样？我准备一会去赛琳娜家。事情挺顺利的。一会儿见，哥们儿。”这里没有深入的分析思考，只是在公众面前保持一个友好的形象。将行动过程中浅浅交流的人与陌生人相比较很有趣，这些陌生人恰好在等待实验或者待在同一个房间里简单地相互认识。像行动中的人一样，陌生人彼此交谈往往是乐观的。不同的是，陌生人更少谈论个人情况，他们的要求更加模糊。

这些相同的语言模式也出现在在线聊天室。试图认识别人的陌生人聊

天室使用的语言像我们的陌生人群体。聊天室里喜爱体育运动的人聊起天来像现实世界中的人在谈论真实的体育赛事。由彼此认识的人组成的聊天室对话和在晚餐时朋友之间的聊天很像。

为了了解这些情境的语言有多么不同，想象一下，我们在计算机里输入几种相似情境下的几百份转录文本。计算机在给语言归类情境方面能做到多好呢？本例中，有五个不同的上下文或情境，意味着仅仅靠猜测，计算机将语言样本分配到正确类别的概率应该达到20%。我们的计算机做得好很多，正确率达84%。

思考一下这些结果的逻辑。所处的情境决定我们使用的语言 。想象一下和你的朋友们一起度过一个周末。如果去看一场体育比赛，你们会使用一套语言模式；来到酒吧，你们会使用一套不同的语言模式。突然间，你发现自己身处一个杂货店，然后又换了另一种语言模式。情境的影响非常强大且可预测，一个人甚至可以说你所处的情境要求你以某种方式说话。好吧，“要求”这词有点过。但是，我们在哪里以及在做什么使我们以非常特定的方式思考和谈话。

表面上看，这可能听起来很平常。运动或观看体育赛事时，人们倾向于谈论比赛。和朋友一起吃饭时，我们大多数人谈论共同好友或过去的经历。情境决定了谈话的主题，这无疑影响着人们隐性词汇或功能词的使用。但还有更多。功能词告诉我们，团体和居住在我们身边的人们关注世界、与他人联系的方式以及思考和感受的方式。

词捕捉也为功能词研究提供了一种有趣的、几乎逆向的思考方式。本书和本章的大部分内容都用来展示人们使用的词汇如何反映他们本身。语言也可以告诉我们人在哪里以及他们在做什么。将对话的抄本输入指定计算机，它会告诉你对话人的信息、关系和情境。

◎探究团体的地理位置

这时事情开始变得有点吓人。如果团体倾向于以类似的方式说话和写作，那么应该能够估计团体的实际位置。这种估计远远超过判断某人是否在商场、在公园或在餐厅吃饭。通过内容词和功能词，我们可以有根据地猜测说话者恰好在哪个商场、公园或餐馆。语言追踪的逻辑与口音识别的逻辑类似。根据你所在的国家，在判断说话者来自哪个地区方面，你可能做得比纯猜测要好。在美国，南方人讲话慢吞吞，中西部人讲话鼻音重，而纽约人的口音常常在几秒钟内就能被听话者识别出来。当然，计算机文本分析程序不分析口音，而是专注于特定的词汇或语言模式，也许这些特定词汇或模式在不同地区是很常见的。

大量的食物、物品和行为的命名也存在地域差异。在美国东北部，如果你想喝软饮，你会要一杯苏打水，在南方你会要一杯可乐（意思是任何类型的软饮，不只是指可口可乐），而在中西部，你会要一杯碳酸饮料。根据你住的地方，你可以管辣调味汁或越南河粉里的绿色叶草叫香菜、芫荽、欧芹或胡荽。如果你不想让前面的卡车把泥土溅在你的挡风玻璃上，你最好希望卡车司机安装上挡泥板（西部）、防溅罩（中西部）或挡溅板（东部）。另外，如果你在英国开车的话，希望车上安装了挡泥板（三个字），在澳大利亚，泥板（两个字）。之后，当你划你的独木舟时，小心别让它倾覆（美国东部和西部）或翻倒（美国南部）。

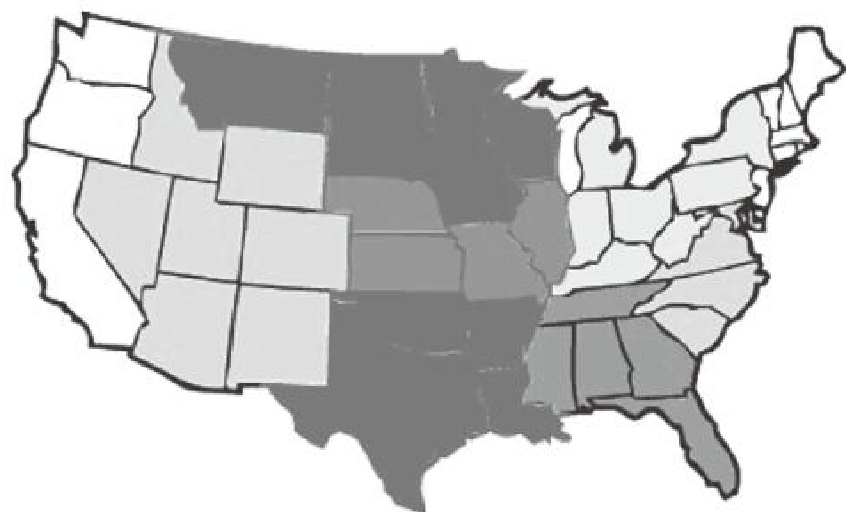
可以想象，通过分类地区命名，我们可以将不同地区的人们分隔开来。剩下只需要等待，等待说话者或写作者提及软饮、香料或他们的船是怎么翻的。与其等待相对模糊的单词或主题出现，不如追踪更常见的词，也许效率更高。也许，功能词？

***

正如你所看到的，隐性词汇根据情境的不同而变化。它们也会根据地理区域的不同而变化。有些地区的功能词使用差异是众所周知的。

即使在相对正式的写作中，功能词的使用也存在地域差异。我和辛迪·钟有机会通过几百篇人们为响应全国性的电台节目“我相信”而写的文章测试这个想法。20世纪50年代，“我相信”的最初版本由记者和无线电评论家爱德华·默罗简单主持。几年中，默罗邀请了当时的一些文化领袖来总结他们最重要的信念。每周广播有政客、体育明星、科学家、哲学家和其他人能够吸引无线电听众的人。2005年，杰伊·艾利森和丹·盖迪曼重拾国家公共广播的想法。不依靠名人，而鼓励听众贡献自己的“我相信”文章。四年的时间里，提交了7万多篇文章，虽然只有大约200篇被广播过。大多数文章最终都公布在网上供公众阅读（www.thisibelieve.org ）。

与艾利森和盖迪曼合作，我们分析了大约37500篇文章。很多故事都很吸引人，有悲伤的故事、有趣的故事、感人的故事和鼓舞人心的故事。区域差异出现在文章本身的主题上。关于体育的故事在美国中西部最常见，南部是种族问题，东北部是科学主题。



a) 直接性示意图



b) 区别性示意图

图　9-1

注：直接性集群中的词（第一人称单数、短词，现在时态动词、无冠词）和区别性集群中的词（排除词、否定词、因果词、无包含性词）的相对使用情况。区域颜色越深表示使用频率越高。语言样本基于37500篇“我相信”文章。

人们的文章不仅主题不同，而且使用的功能词也不同。正如你在直接性示意图中可以看到的，美国中部的人们倾向于使用高频率的“我”、现在时态动词和短词。回忆一下之前的章节，这个词群反映了心理上的直接性，写作者往往处在当下。图9-1中，较暗的区域反映了更高的直接性。来自东北部和西部的写作者的文章个人情感最少，社交性最低，且最具体。这些低直接性的文章反映了心理疏离和正式的语言风格。

回想前面的章节，反映分析性思维方式的功能词常聚集在一起使用。具有分析性思维的人常常在想法之间进行区分。做区分时，有必要使用如连词（但是、如果、或）、否定词（没有、不）和介词（和……一起、在……之上）这样的词。在区别性示意图中，美国中部的人最常作区分，而东北部的人作区分是最少的。

为什么会存在这些区域差异？正如从语言风格匹配上看到的，人们很快地调整他们的说话风格以适应周围的人。交谈的时间越长，以相似方式开始看世界的人越多。一般情况下，我和我的邻居享有相同的天气、吃类似的食物、参与相同的社区活动，并且学校、税收、商店和政府体制都一样。我的社区的人和隔壁城镇的人共享很多相同的事情，在一定程度上与相邻州的人共享很多相同的事情。但是，当我离家乡越来越远时，天气、食物、文化和关心的事情就发生改变了。随着社会环境和物质环境的变化，人们接近世界和与他人交谈的方式也在变化。

***

虽然随着距离的变远，语言差异应该更加明显，但是有些语言差异也会产生在天气、地形、种族、社会等级和其他每个因素都相似的很近的地方。我在西得克萨斯州一个盛产石油的小镇长大，很多家庭在搬到其他地方前会在这里生活大约四年。即使不断有人迁移至此，新来的孩子很快就会适应当地的口音和俚语。

甚至在学校里，老师能够区分不同的小圈子里不同的语言模式。20世纪80年代初，对底特律一所高中的重要分析中，佩内洛普·埃克特证明，该校的校园语言与众不同。与世界上大多数中学不同，该校形成一个紧密连接的团体，采用他们自己的语言风格，以此反映出他们的团体身份。

不难想象，在同一地理区域的不同学校可以发展自己的语言风格。通

过分析5万多份学生提交的入学申请文章，我们找到了这方面的证据，这些文章是多年来得克萨斯大学奥斯汀分校所接收的。与学校的招生办公室合作，我和语言学家大卫·毕佛看了来自得克萨斯州的一个大市区的九所不同高中的2000篇文章。来自各个高中的学生在高中和大学第一年期间表现同样良好，社会阶层和种族构成稍有不同。然而，他们在他们的文章中使用代词、冠词、介词和其他功能词的方式因学校而异。换句话说，每所学校都有自己的语言指纹。

和“我相信”故事一样，大学入学申请文章写作形式独特。大多数人一生只会有几次写像自己的文章——如果有的话。通常，作者单独坐在他们的房间（或咖啡店），思考他们生活中一些更大的问题。他们的故事反映了他们的家庭、朋友、社区，以及社会。指导他们语言选择的内心声音是他们的想法、他们关心的事情、他们的情绪状态和他们的语言历史。

毫无疑问，自我反思的文章反映了人们的地域感和他们长时间相处的团体。即使通过相对简单的计算机模型，我们也可以通过文章中使用功能词的方式来估计一个人在哪个国家的哪个城市，甚至在城市的哪个区域。如果我们分析人们对话中的文本，同样的功能词会提供关于人们正在做的事情、他们所处情况的线索，以及他们与周围人的联系性质。

如果你有妄想倾向，要知道，基于功能词的掠夺者无人机是不可能被开发的。我们的语言总是反映我们是谁、我们在哪里以及我们在做什么。是谁、在哪里、做什么是一直很明显的。在写文字之前，如果我和你说话，我们都知道我在说话（谁）、我们的位置（在哪里）和我们当前的行动（什么）。多亏技术的进步，我们才有了一个时期，使得我们是谁、在哪里和做什么变得不透明。有趣的讽刺是，随着技术的进步，我们可能最终能够还原5000多年前我们的祖先是谁、在哪里、做什么。

***

本章探讨了人们在团体中使用的词语如何揭示团体本身的一些东西。团体成员使用“我们”通常表明成员认同团体。一段时间后，随着人们在团体中变得更加舒适，每个人都倾向于使用更多的“我们”。当群体成功或受到外部威胁时，群体身份增加，相应地增加使用我们的词语。

“我们”的使用反映团体认同，但不能反映团体在一起工作的和谐程度。提升一个团队的认同感是可能的，但这并不意味着团队真的会做得更好。可能团队中没有小“我”，但团队中也没有“我们”。语言分析表明，团队成员为了达到最好的工作状态，必须以相似的方式思考，并密切关注团队其他成员。十分可能的是，语言风格匹配反映了团体中不同成员之间的相互喜欢和尊重。

在本章中，团体的定义相当宽松。这是可以的：我是一个社会心理学家。特别吸引人的是，在约会的情侣、工作实验室团体、现实世界工作组、在线社区以及整个学校、城市和社会中，“我们”的使用过程相似，且语言风格匹配显而易见。统一的主题是所有这些团体都使用语言来沟通。无论是书面形式还是口头形式，语言是互动的共同媒介。

最后，就像团体成员的语言揭示团体的信息一样，它们也揭示了团体成员正在做什么以及他们在哪里的信息。奇怪的是，功能词使用具有高度感染性。无论是情侣、小团体、社区还是城市，人们都倾向于采用他们周围人的语言风格。我们的语言，特别是我们使用的功能词，无意中揭示了我们正在做什么以及我们在哪里。正如我们的口音、肢体语言和穿着揭示了社会自我和心理自我，我们的语言也一样。

如果你是一个私家侦探，收起你的小望远镜。取而代之，打开你的电脑开始数词吧。

第10章　语言侦察

功能词的作者识别可以帮助抓捕罪犯或确定历史作者，可以帮助了解总统或暴君的思想，可以帮助预测人们未来的行为，还可以开始更好地了解我们自己。

在学习语言时，我经常被要求通过分析语言来回答我从来没有考虑过的问题。律师、历史学家、音乐爱好者、政治顾问、教育家、情报员和其他人偶尔会联系我，看看我们的语言方法能否在他们一直在思考的问题上给出不同观点。本章汇集了我和我的学生多年来一直在考虑的一些更有趣的项目。项目主题很多，它们展示出可以回答新奇问题的不同语言分析方法。

通过语言识别作者

一家律师事务所的高级合伙人打来的电话让我措手不及。他很好奇我能否分析发给他们公司成员的一封电子邮件，我们暂且称她为利文斯顿女士。他坦白地说，这是相当敏感的事情，而且重要的是，他直接与发送电子邮件的人交流。唯一的问题是，电子邮件是从一个不可追踪的电子邮件地址匿名发送的。在我同意看一下邮件后，他发给我下面这样一封邮件。

利文斯顿女士：

我想你应该知道，大卫·辛普森坚信你在你的同事中可信度为零。他说你在纽约的最后一份工作中改变过证词、伪造费用报表。他说这是你突然离开的原因。

他将这些故事传播给了各部门的人，包括财务部、人事部、公关部以及领导层的人。我们的高级合伙人将如何以及何时处理这一问题还不确定。但如果你开始受到冷落，你知道是为什么。

当我第一次听到这个时我很惊讶，但表面上接受了他所说的。当然，这是在我了解到他对传播半真话、八卦和彻头彻尾的谎言的贪婪欲望之前，所有这些都是他在以某种方式使自己看起来见识广和“更优秀”。

真是遗憾。显然，他很有天赋，但都被他的恶毒和邪恶所毁掉。我所能想到的是极其缺乏安全感。离题了。我只是觉得你应该知道这些。

一个朋友

在收到邮件后，利文斯顿女士把它转发给了律师事务所。她证明了传闻是假的，但是她关心大卫·辛普森是否真的在散播谣言，这可能会损害她和公司的声誉。我花了几年时间研发分析语言和性格的方法，但从来没有人付钱给我做语言侦探。

什么样的人会写这么一封邮件呢？“一个朋友”是男性还是女性，他或她的年龄大概是多少呢？这个人与利文斯顿女士、大卫·辛普森以及公司是什么关系？关于这个人的性格特征有何线索？

在我从事这项工作以来的几年里，已经开发了几种研究文字的新方法。其中一种方法涉及将“一个朋友”与成千上万个常规博主的用词进行比较。例如，仅通过观察功能词和情感词的使用，我们就可以猜测，作者是女性的概率为71%，她年龄为35～45岁的概率为75%。要通过阅读更好地

了解她的性格，要难得多。一个分析表明，电子邮件的作者很有可能有自恋特质，这意味着她可能有点自负和爱指使人。

仔细阅读电子邮件，出现了其他线索。这个人在心理上与公司（“我们的高级合伙人”）很近，并且听说了来自公司内部好几个部门的谣言。这个人也很努力地通过使用大量词汇来打动利文斯通女士。特别有趣的是使用“贪婪欲望”“恶毒”和“邪恶”这样的词。这些是《旧约》中的词，其他分析表明，在当时，使用者主要是42～44岁的人。

另一条重要的线索是页面布局和标点符号。电子邮件的段落同等宽度，页面排版很专业。句与句之间只有一个空格，表明发邮件者是在1985年后学习的打字（那是台式电脑开始流行时）或者发邮件者在1985年之前有新闻或出版背景，因为在这两个行业里，句号之后一个空格是标准。（我的妻子1985年之前在出版行业工作，她这样向我解释。）

发生了什么？当我向那位高级合伙人提交我的报告时，他感到如释重负，因为我的判断与他怀疑的人精确匹配：40多岁具有新闻背景的一位正直的女性，她在公司已经工作了好几年。我不知道事情的最终处置情况，但我知道，利文斯顿女士现在是该公司的一位高级合伙人。

谁写的？作者的身份识别

破译语言线索来解决犯罪问题在刑事学中是一种惯例。美国联邦调查局、各国家安全机构和世界各地的地方警察部门偶尔会寻求语言专家的专业知识，来帮助解释勒索信，评估法律文件或其他文件由谁执笔完成。

最著名的早期法律语言学家之一是瓦萨学院英语教授唐纳德·福斯特。运用计算机和推论技巧相结合，加上历史和文学知识，福斯特与执法机构合作，一起攻破备受瞩目的案件，例如“智能炸弹客”、2001年的炭疽攻击和1997年的乔恩·贝尼特·拉姆齐谋杀案 [1] 。他还运用他的方法来鉴别莎士比亚和其他人一些作品的真伪。也许他最成功的事迹是确定乔·克莱就是匿名发表讽刺克林顿总统小说的作者，书的名字是《风起云涌》（Primary Colors）。

福斯特一直是一个有争议的人物，因为他有些备受关注的关于作者的判断没有成功。而且，他不愿意透漏作者身份识别方法的细节，这些方法反映出他的英语很好，而不是根据统计数值和科学。尽管如此，福斯特的方法已经为文学界和法律界用以计算机为基础的方法来识别作者及其作品的未来打开了大门。

◎找出破绽

世界级扑克玩家会密切观察对手和听对手发言，试图猜测他们手中拿的什么牌。通常当玩家有一手好牌时，他们会假装牌不好；而其他时候，他们会虚张声势，做出一副赢定了的姿态。老手会寻找说谎或露出破绽的迹象。一些玩家避免四处观望，有些人轻轻跺脚，有些人会说话更大声。玩家具有识别破绽的能力在高赌注扑克游戏中有很大的优势。

人们使用的书面语言中也有各种类型的破绽。在识别作者方面，有两个特别好的线索：功能词和标点符号。这可以从我们在2001年收集的博文中看出，这些博文是上一章讨论的“9·11”项目的一部分。回忆一下，我们保存了大约1000人的博文，每人平均70条，发自“9·11”事件前后两个月。每隔几年，我和我的学生都会再次访问LiveJournal.com，看看这些用户是否仍然发布博文。十年后，25%～30%的用户仍然活跃在网站上。大约25%的用户注销了他们的账户。2006年，事件发生后的5年，其余的用户相继停止发布博文。许多以前的帖子迁移到其他系统上，如脸书或推特网。

只是阅读过去十年人们发的帖子，就可以对他们的生活有一个清晰的画面。不同于迈克尔·艾普特的7 Up纪录片系列，随着博客用户的年龄增

长，我们追踪了他们的人生经历。许多同样的问题仍然驱使着作者。即使有些人已经结婚、生子、创业，但是重复的不安全感、动机和目标不断复现。那些在2001年快乐和乐观的人往往在九年后同样是个乐观的人。例如，一个年轻的父亲在2001年的博客中随手写到他最喜欢的曲棍球队。

幸运的鸡骨。我应该跳“快乐杯舞”。我们会赢。我们将胜利。会很开心！我只需要有一个台。好。我不只是要发布关于曲棍球的消息，但伊凡已经准备好了。是的。闭嘴。你试着控制住不看电视。

九年后，你会看到同一个人。

我第一次做辣调味汁，依我看，没那么不堪。专业意见：不要使用罗马牌番茄酱。我不知道为什么我会认为它们会很好吃，但显然我彻底错了。好吧，不是彻底，只是有点儿。哈，辣酱式幽默。我今天要去墨西哥超市买东西，然后再尝试一次。也许我会在烧烤开始前及时完成。谁知道呢？自从我发泄完，我现在脑洞大开。

显然，这两个样本来自同一个人。我的意思是，任何人都可以感觉出来。

真的吗？

实际上，一旦我们知道他们是同一个人写的，我们就可以看出相似之处。但是，如果我们一整天都在刷博客，看完第一个博文几个小时后才读到第二个呢？十有八九，大多数人不会跳起来喊：“啊！我之前读到过类似风格的……没错，是那个描述曲棍球的人。”语言专家或计算机可以做出明确的匹配吗？语言指纹是否像DNA或真实指纹那样可靠？简短的答案是不。但是，计算机语言分析在风格和作者匹配方面做得挺好。

想象一下，我们有20个人写的大量博文。几年后，我们从同样的20个人的博客中检索出一小部分新的博文。现在想象你坐在客厅地板上，眼前有数百页的博文，你试图将当前的每条博文与20个博主之前的博文进行匹配。所有条件都是均等的，即使只是靠纯猜测，每个人匹配的正确率都应该能够达到5%。大多数人会做得很差。匹配正确率多处于10%～12%，都不太可能高于这个数。写作风格的差异太微妙，而且资料太多。

计算机更有耐心和更有条理。如果我们只是分析功能词，计算机正确匹配的比例在当时大约29%。这实际上是令人印象深刻的，因为博文和博文之间存在时间差。

但作者识别不仅仅只是功能词。看看标点符号的一致性。例如，下面

这个女性，相隔9年，使用星号的方式仍然相同。这是2001年某条博文的一部分：

噢……我还发现了我不曾知道我有的害羞倾向。我猜你会称之为羞怯。某人使我*脸红*。一次又一次地。这很*奇怪*。我从不会脸红。

2010年：

我们现在*处于*后后朋克时代了，不是吗？乐队的伙伴开了一个关于昨天刚写的新歌的玩笑，可能屋里有1/4的人都没有懂我们为什么笑。奇怪。*耸肩*

其他人使用标点符号的方式同样独特但更微妙。一名27的男性写于2001年：

我寄了纪念礼物支票给以马内利（捐赠基金以纪念琼的母亲）；和圣·安——为我最爱的会计学教授，斯迈思奖学金。弗兰克＆丽贝卡带来了《午夜善恶花园》和夫妻自制啤酒。我的眼皮要睁不开了，所以我最好……

2010年：

我真不知道说些什么，想着“嗯”，泥，那是什么……当我找到一面镜子我没有看见任何“棕色的东西”我带了一个西瓜和好市多杂粮薯片，喝了几瓶酒，我喝了云岭B&T——晚饭是煮/烤鸡、秋葵、卷心菜沙拉、“涂层”布朗尼。

这个人简直是标点符号舞蹈家。他跳跃、旋转、飞扑和翻转，具有全方位的标点符号可能性：[;-...&“/。奇怪的是，当我第一次阅读他的博客时，我都没有注意到他标点符号的使用，因为它们自然地融入他的写作中。然而，当计算机分析他的博文时，他的标点符号使用特点显现了出来。

标点符号对于一些人来说比他们写的任何东西都更好识别。事实上，仅通过标点符号，计算机识别作者的正确率为31%，基本上与功能词识别比例相同。当功能字和标点符号一起使用时，计算机配对原始博文和几年后博文的正确率在当时是39%。

邦党人文集

1787和1788年，一系列共85篇文章被出版在小册子和报纸上，横跨美洲殖民地，试图影响人们来支持将成为美国宪法的文件。文章匿名发表

在普布利乌斯（Pabulius）名下，文章讨论的主题非常广泛，包括总统角色、税收、州政府和联邦政府权力等。即使在当时，许多人知道普布利乌斯不是一个人，而是詹姆斯·麦迪逊（将成为第四任总统的人）、亚历山大·汉密尔顿（财政部第一任部长）和约翰·杰伊（最高法院的第一位首席法官）。

随后几年里，74篇文章的作者逐渐被人们所知晓。麦迪逊写了15篇，汉密尔顿51篇，杰伊5篇，麦迪逊和汉密尔顿联合撰写了3篇。剩下11篇文章的作者一直都没有确定，自那以后成为推测的来源。第一个认真寻找11篇文章作者的人是历史学家道格拉斯·亚岱尔，作为他1943年一篇论文的一部分。亚岱尔的历史分析推断，所有11篇匿名文章都是由詹姆斯·麦迪逊所写。

1964年，当统计学家弗雷德里克·莫斯特勒和大卫·华莱士引入一种新的分析方法时，争论重新出现。通过关注少量的功能词，他们得出结论，亚岱尔确实是正确的，因为他们精密的统计模型指出麦迪逊可能是作者。从那时起，每当开发出新的语言分析方法，确定联邦党人文集的作者就成为一种娱乐项目。

我自豪地宣布新的官方调查结果。历史学家，准备好你们的鹅毛笔。

功能词分析

通过与莫斯特勒和华莱士相似的方法，我们得出同样的结论。匿名的11篇文章使用代词、介词和其他功能词的方式都与詹姆斯·麦迪逊相似。结案？

没那么快。在对我们开国元勋的功能词调查中，其他统计学家发现了一个小问题。自从莫斯特勒和华莱士后，另一种被称为交叉验证的技术出现。这个想法是单独检查每篇原始的文章，就好像它们是匿名写的。换句话说，我们从文章堆中提取出一篇已知的文章，然后基于剩下的文章开发一个计算机模型，以试图确定我们提取出的文章出自谁手。这是一个奇妙的方法，因为我们在确定一个已经知晓答案的问题的结果。如果交叉验证分析成功地猜出谁写的这些已知作者的文章，我们就可以充分信任我们的研究方法。

伤心。交叉验证结果表明，莫斯特勒和华莱士可能是错的。基于功能词使用，大约14%已知作者的文章没有被正确分类。这是一个严重的问题。如果计算机不能以非常高的准确度告诉我们已知的结果，那我们对于作者未知的文章分析结果就要小心地解释。

标点符号分析

回想一下，许多情况下，人们使用的标点符号可以揭示作者的身份。使用类似的交叉验证分析标点符号，结果也令人失望。使用功能词和标点符号组合来预测作者身份比单独使用功能词的结果稍微好一些。有趣的是，功能词加标点符号的结果显示，汉密尔顿写了11篇匿名文章中的3篇。

找出破绽：模糊词汇的使用

在我的职业生涯中，我写了很多文章。大概10年前，一位同事感谢我对她的研究做出的评论。当然，我感到荣幸，但有点困惑，因为我的评论是匿名写的。“你怎么知道评论是我写的？”我脱口而出。她笑着只说了一个词：“引人入胜”。

引人入胜，的确。我翻看了很多我的评论、文章甚至书。我使用“引人入胜”这个词的频繁程度让我感到震惊。甚至本书也是，因为我就是控制不住自己。多年来我注意到，我的大多数同事和朋友都有自己喜欢但相对模糊的词，虽然他们也不熟悉。这些词的使用频率不是很高，但偶尔在电子邮件、脸书、博客、推特或文章中能够找到它们。

麦迪逊和汉密尔顿的文章中有这样的词吗？通过小小的侦查，发现答案是肯定的。在他一半的文章中，汉密尔顿都用到“容易地”这个词，麦迪逊从来没有用过。15篇文章的9篇中，麦迪逊都用到“因此”，而汉密尔顿在他的51篇文章中只用到过3次这个词。汉密尔顿还喜欢使用“通常地”“足够的”“有意的”“善良的”和“自然地”。麦迪逊倾向于过度使用“绝对地”“管理”“背叛”“组成”“包围”“创新”“谎言”“会议记录”和“愿望”。

如果我们只检查这14个词的使用，数据是很明朗的——对于交叉验证来说，几乎是完美的数据。然而，未知作者的结果与先前学者的判断相差很大。它们显示，汉密尔顿写了8篇匿名文章，麦迪逊只写了3篇。

实际发生了什么？阅读道格拉斯·亚岱尔围绕这11篇文章的争议写的有意思的说明，很明显，汉密尔顿和麦迪逊对于谁写了哪些文章有着不同的回忆。亚岱尔最终更加赞同麦迪逊的说法，尽管客观证据并不是很能令人信服。同莫斯特勒和华莱士一样，我对实际情况没有深入的了解。然而，历史学家应该知道，从统计的角度来看，本案仍没有了结。

我们每天都在使用标点符号、功能词和内容词，它们是我们个性签名的一部分。为了更好地理解这一点，请花几分钟时间查看一下你发给别人和从别人那里收到的电子邮件。从页面布局开始。有些人倾向于写很长的

电子邮件，有些人则一两句话解决。每个人写的段落和句子长度往往不同。问候句和结语差距也很大。有些人使用表情符号，有些人从来不用表情符号。

从心理角度讲，有些差异可能很重要，但多数不是。有些人结束语多用“祝好”，可能只是因为小时候他们被告知要这样做。虽然这些变化可能不会说明你婴儿时期与母亲的冲突，但是它们仍然是你的标记。也就是说，它们是你写作风格的一部分，使你与其他人有所不同。这就是故事有趣的部分。所有可以测量的语言特征都可以帮助我们识别你。

[1] 乔恩·贝尼特·拉姆齐（JonBenét Ramsey）是一个六岁的选美大赛王后，1996年12月25日在家中被杀。——译者注

歌词揭示乐队情况：披头士乐队

披头士在1970年解散前在一起组合了十年。组合期间，他们录的歌曲超过200首，影响了音乐界、政界、时尚界和下一代的文化。主要作曲者约翰·列侬和保罗·麦卡特尼一起或分别写了155首歌，乔治·哈里森又写了另外25首。即使在今天，学者和偶尔去酒吧的人仍在讨论乐队成员的相对创造力，谁最终影响了谁，以及乐队如何随着时间发生变化。

本书的大部分内容都是关于人们在谈话、文章、信件或电子媒介如博客、电子邮件中产生的话语，然而，音乐的歌词也讲述着它们作者的故事。我的新西兰好朋友兼合作者基思·皮特里提出，披头士的计算机语言分析早就应该进行了。当我们意识到这个话题真的很复杂时，我们邀请了挪威的另一位音乐爱好者兼心理学家博奇·西尔森加入我们。我们可以通过分析披头士乐队的歌词可以了解到什么？结果，了解到很多。

在许多方面，乐队的歌词反映了所有工作团体中通常会看到的自然衰老过程。回想上一章，随着工作团体在一起的时间变长，他们对话中的“我”字减少，“我们”增多，语言复杂性增强，包括更大的词和更多的介词、冠词和连词。随着乐队成员一起变老，就像一个团队在彼此的对话中一样，披头士们在他们的歌词中以同样的方式表现出来。

在他们组合的最初四年里，他们的歌曲充满了优雅、愤怒和感性。他们的想法很简单，专注于自我，活在当下。乐队最后几年里，歌词变得更加复杂，在心理上更加遥远，并且很不积极向上。特别有说服力的是，乐队第一年的“我”字使用频率近14%，最后三年下降到只有7%。歌词还为乐队成员的性格提供了一个窗口。虽然约翰·列侬和保罗·麦卡特尼有一个协议，他们所有的歌曲都由两个人共同作曲，但是作者的顺序和广泛的访谈为谁是作者提供了一个虽然不完美但可靠的记录。两人之间，列侬被认为是78首歌曲的主要作者，麦卡特尼67首，另外15首被认为是两人都密切参与创作的真正合作。

流行媒体普遍把约翰·列侬描述为有创造性的文化人，而麦卡特尼是音调优美、欢快的流行歌曲作曲者。对他们歌词的分析描绘了不同的画面。列侬在歌曲中使用的消极情绪词汇确实比麦卡特尼稍微多一些，但是两者在积极情绪、语言复杂性和自我反思词汇的使用上几乎相同。有趣的是，麦卡特尼的歌曲往往侧重于情侣，这可以从他使用的“我们”中看出——高于列侬。

谁更富创造力或者说谁的歌词创作更具多样性？我们可以通过不同歌

曲的歌词在数值上的相似度来测试这一点，无论是在内容还是语言风格方面。虽然流行媒介通常认为列侬是创造性和风格多变的作词者，但是显然数字支持麦卡特尼。在作为披头士一员的乐队生涯中，保罗·麦卡特尼证明，他在写作风格和歌词内容上都更加灵活和多样。

不要忘记还有乔治·哈里森，那个安静的、精神的披头士，他写了大概25首歌，特别是在披头士乐队的最后几年。虽然他的语言比列侬和麦卡特尼在认知上更具复杂性，但是他的写作风格最不具灵活性。换句话说，他的每首歌的歌词内容和风格可预测性更强。这些相同类型的分析也表明，哈里森在作品风格上受到列侬的影响比受到麦卡特尼的影响更多。

小结：整理你的作者识别工具包

作者识别逐渐成为计算机世界中的一个非常热门的话题。我们所依赖的三种方法涉及追踪功能词使用比率，分析标点符号和布局，以及检查模糊词汇的使用。这些方法中的每一种在识别作者特征以及将作者和作品匹配上都比纯猜测做得好很多。

合作产生平均结果还是协同效应

作家之间的合作是一件有趣的事情。当两个人一起工作时，用约翰·列侬的话说就是“面对面”，产生的作品是他们通常风格的平均水平，还是完全不同于任何人单独写作可以完成的结果？语言分析可以为披头士乐队和联邦党人文集回答这个问题。回想一下，列侬和麦卡特尼在他们160首歌曲中的15首上有非常密切的合作。亚历山大·汉密尔顿和詹姆斯·麦迪逊共同撰写了三篇联邦党人文章。

从语言的各个维度甚至是标点符号上，我们可以计算出何种比例的合作产生的效果是两个协作者单独写作的平均效果。有三个明确的假设如下所示。

·像合作中的一个成员假设。 协作性写作项目产生的语言类似于单独一人的写作。作品有时像其中一个作者，有时像另外一个作者。

·平均假设。 更有趣的是，合作产生的语言是两个作家的平均水平。如果列侬使用“我们”的比例较低，而麦卡特尼使用的比例较高，那么他们的合作会产生中等数量的“我们”。

·协同效应假设。 更有趣的是，当两个人合作时，他们创造的作品不同于任何一方。他们的语言风格将具有独特性，使得大多数人不会意识到作者是谁。如果分析结果支持这个假设，那不是很好吗？拜托，数据，请让我开心。

事实上，协同效应假设胜出。当列侬和麦卡特尼合作，以及当麦迪逊和汉密尔顿合作时，他们的作品与每个作家单独的作品截然不同。合作时，列侬-麦卡特尼团队制作的歌词比任何一方自己的歌词都更积极，同时使用更多的“我”、更少的“我们”和更短的单词。同样，当汉密尔顿和麦迪逊合作时，他们使用了比自己作品更大的词、更多过去时态和更少的助动词。事实上，在语言和标点符号的大约75个维度中，超过90%的维度在合作语言中高于或低于两个作者自己的语言。

请注意，合作产生的语言模式与个人单独自然产生的语言模式完全不同。现在还不知道的是，合作作品是否总体上比个人作品更好。这是一个急待研究的问题。

在了解作者性格方面，我们目前研究最多的是功能词。正如本书所讨论的，代词、冠词和其他隐性词汇确实与作者的年龄、性别、社会阶层、性格和社会关系相关联。目前关于标点和性格的了解较少，但我猜想未来的研究会证明出令人信服的联系。毕竟，很难想象在纸条结尾处写着“谢谢”和“谢谢！！！！！！！”的两个人之间没有任何区别。

了解最少的是相对模糊词汇的使用及其与性格的联系。如果一个作者使用“引人入胜的”，另一个使用“非凡的”，这个词本身的选择说明了什么吗？我呼吁未来的科学家回答一下这个问题。

还有许多与作者识别相关的令人兴奋的方法，由世界各地的实验室开发。其中一个方法是观察一些称为多元的东西。这些可以是二字词（或二元），三字词（或三元）等。这个想法是，有些人自然地以独特的方式使用词组，通过这个可以识别出他们是谁。

更复杂的方法试图基于作者已经使用的词汇，在数学上预测句子中的词序。上一段的开头以“这”开始的可能性也许是1/1000。知道第一个词是“这”的情况下，第二个词是“是”的概率也许是1/20。知道前两个词是“这是”的情况下，第三个词是“个”的概率是……明白了吧。研究人员可以确定一个人的写作有多特别，以及它在句子级别上偏离的概率。一个论点是，每个人将词汇串在一起的方式都是特有的。这也是另一种语言指纹的想法。

其他新方法会检查词性、语法、句子和段落的连接性——所有都使用越来越复杂的数学方法。不久的将来，大多数展开的语言样本的作者都是可识别的。

政治事件和历史事件的语言线索

对历史学家和文学学者来说，对书面文字的研究是了解人们或过往作品的主要关键，这并不是什么新闻。然而，大多数学者主要依赖于他们自己阅读的历史著作，而不是计算机化的文本分析。这在过去几年一直在变化。一个特别创新的领域是政治学。研究人员能够探索政治候选人的吸引力和人们对他们的反应，部分原因是演讲、采访、报纸和在线文章、新闻广播，甚至给编辑的信件的转录文本的可用性。

该领域的先驱之一罗德里克·哈特已经出版了一系列突破性的书籍和文章，这些作品有助于解释重要的历史选举结果是如何通过候选人演讲中的词汇使用方式预测出来的，如比尔·克林顿和鲍勃·多尔之间的竞选。他还收集了数百封寄给美国各地新闻编辑的信件，并能够追踪选民的看法。延伸哈特的开拓性工作，我们可以通过分析所有留下语言痕迹的历史人物的语言，重新解释历史事件。

这些人是谁？尝试了解美国总统

观看任何新闻（可能来自电视、互联网、报纸、杂志），其中大部分内容都是为了了解当前、未来和过去的总统的想法。如果在一场竞选的中间阶段，专家们会推测每个候选人将如何参与竞选。如果一位总统最近被选举或重选，我们想知道其打算在未来几个月和几年内完成什么。即使在总统下台后，专家们仍然会问："他在想什么？他为什么这样做？"

在政治心理学最令人印象深刻的著作之一《政治头脑》（The Political Brain）中，研究员德鲁·韦斯滕认为，最成功的政治家是能够与选民进行情感联系的人。逻辑、智力和理性当然是非常优良的品质，但当选民参与投票时，通常推动选举进行的是社会和情感因素。

我们会与那些看起来关心他人、尊重他人，同时真实表达自己情感的人产生共鸣。我们可以通过身体语言、语调，当然，也可以通过语言来探测人的社会-情感风格。对于总统和总统候选人，我们有充分的机会通过演讲、访谈、图片以及他们和家人、其他人的互动视频来评估他们的社会-情感风格。从语言的角度来看，总统留下了大量不同于其他人的语言。

测量社会-情感风格的一种相当简单的方法是计算使用人称代词和情绪词汇的频率。一般来说，自我反省和对他人感兴趣的人使用人称代词的频率较高，包括"我""我们""你""她"和"他们"。同样，与不使用情绪词汇相比，如果人们使用情绪词汇（包括积极情绪词汇和消极情绪词汇），那么他们情感在线。通过分析总统演讲中的代词和情绪词汇，我们可以开始了解他们通常情况下的社会-情感风格。

美国总统最多每四年发表一次就职演说。但是在职期间，大多数总统每年都向国会提交国情咨文。国情咨文开始于乔治·华盛顿，1790年。华盛顿和约翰·亚当斯以演讲形式呈现给国会，而到托马斯·杰斐逊时改变了这一传统，只是简单地提交了一个书面版本。1913年，伍德罗·威尔逊恢复了演讲形式的国情咨文。然而，1924～1932年，国情咨文回到书面形式。从1933年富兰克林·德拉诺·罗斯福（简称罗斯福）的就职典礼到今天几乎所有国情咨文，都以演讲形式提交给国会。尽管呈现形式不同，但是观察总统和总统之间情绪基调的变化也令人着迷。

从图10-1中可以看出，几位总统的社会-情感比他们的前辈丰富得多。托马斯·杰斐逊、安德鲁·杰克逊、西奥多·罗斯福、吉米·卡特和乔治W.布什使用人称代词和情绪词汇的比例很高。证据显示，詹姆斯·门罗、沃伦·哈丁和巴拉克·奥巴马的社会-情感评级远低于他们的前辈。事实上，很有趣

的是，乔治W.布什无疑是美国政府历史上最具社会-情感的总统，奥巴马目前是继理查德·尼克松以来社会-情感最低的人。

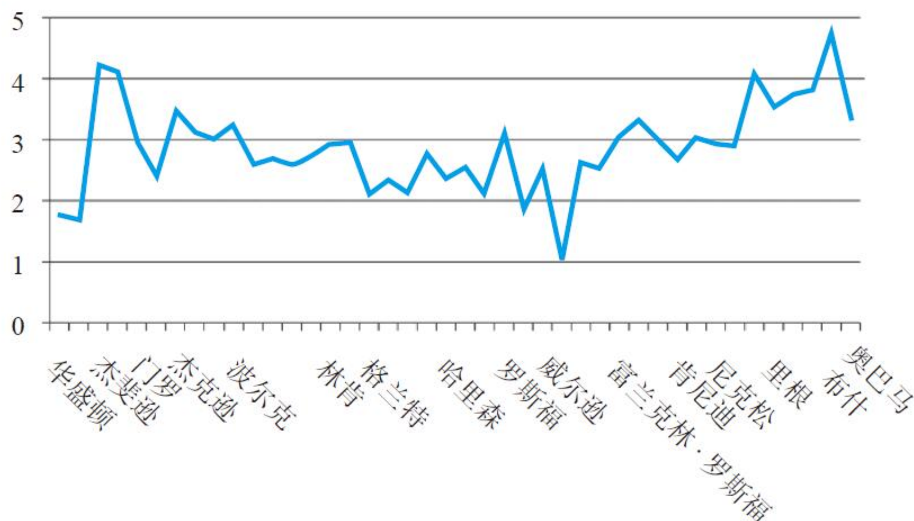


图10-1　社会-情感风格：1790～2010年国情咨文

注：美国总统在向国会提交的年度国情咨文中社会-情感语言的比率。数字已根据书面和口头形式的不同而调整

国情咨文本质上是一个正式的讲话，理论上可以由任何人执笔。它的基调可能反映了整个管理层，但它不一定告诉我们总统本人的心理结构。幸运的是，由于新闻发布会的普及，如今总统的语言有一个更自然的来源。从罗斯福开始，新闻发布会演变成媒体和总统之间的自由互动，并被转录和保存。

从心理学家的角度来看，新闻发布会非常好。媒体成员以各种方式试图哄骗、扶助、挑战、激怒总统。媒体和总统的关系更复杂，因为媒体和总统都需要彼此完成他们稍有不同的目标。最重要的是，与新闻媒体的互动通常是无剧本的，我们能够通过总统使用的语言来持续观察其想法和情绪。

大多数总统每年的新闻发布会很少，通常为4～10次，如果有国家危机，次数会增多，如果被公众嘲笑或被弹劾，则次数减少。除了正式的新闻发布会，总统经常会在随机场合与记者交谈，例如在介绍完一位外国官员或在等待他们的专车时。因为大多数与新闻媒体的会议都会被记录和转录，所以自20世纪30年代以来，大多数总统的自然语言样本通常会有几十

个。有趣的是，总统与媒体的谈话方式在语言上与他们的国情咨文很相似。和国情咨文一样，乔治W.布什的语言在新闻发布会上是最具社会-情感的。自罗斯福之后，尼克松是到他为止社会-情感度最低的总统。

罗纳德·里根呢？许多人认为里根是自罗斯福以来最擅长社交的总统之一。在演讲和新闻发布会上，里根使用的人称代词和情绪词汇总是处在平均水平。外界来看，他似乎是一个富有社会-情感的人——有点儿像乔治W.布什。但是对里根语言更深入的分析表明，这可能是一个错觉。里根看起来似乎更像是一个讲故事的人，而不是一个富有社会情感的领导者。前面的章节提到，讲故事或叙事需要使用社交性词汇和过去时态的动词。结合这两个维度，里根的讲故事评分远远高于任何其他现代的总统。

罗纳德·里根的研究结果让我们更深入地了解那些可能擅于社交或品貌兼优的人。无论政治观点如何，大多数与乔治W.布什接触过的人都觉得他的社交能力很强。在社交场合，他对其他人很感兴趣，并且很容易表达自己的情感。不论是真是假，大多数人走的时候，都觉得自己了解了他。

里根的传记描绘了一幅完全不同的场景。事实上，里根的官方传记作家埃德蒙·莫里斯最终放弃了写传记这一传统，因为他不能让里根打开心扉。里根喜欢讲各种故事，但据莫里斯说，他“对其他人没有兴趣”。在做过1998年关于里根的深入的上下系列电视节目之后，系列编辑阿德利亚那·博世报告说：“里根不是一个惯于自我反省的人……他的儿子罗恩告诉我们，‘没有人知道他在想什么，他自己都不了解自己’。”

虽然外人可能天真地认为布什和里根都是富有社会-情感的人，但是语言分析有助于探索这些印象之外的本质。

“我”说了什么？——奥巴马少用的人称代词

有人希望你在阅读本书时一直在做笔记。如果你有在做笔记，请参考第一人称单数“我”的使用方式。也许你跳过或已忘记了前面章节的内容，但觉得自己可以做一些“我”字使用的跳级测验。你可以的。请访问下面的网站并参加一分钟的测试：www.SecretLifeOfPronouns.com/itest 。对于那些有笔记的人来说，这也许也是一个好主意。

完成关于“我”的10项测试的人现在已经超过2000个，结果显示很少有人知道谁更常使用“我”字。事实上，语言学博士和高中毕业生在这一测试上是一样的，平均正确率都是5/10。如果你做得不好，你有很多同伴。

“我”是典型的隐性词汇。口语中很常用，但是我们很少注意到我们或其他人在使用它。因为人们认为“我”反映出一个人的自信或傲慢，所以他们认为自信的人一定一直在使用“我”字。

奥巴马是一个完美的案例研究。2008年当选后的几天内，专家，尤其是那些不支持他的人开始注意到他一直使用“我”字。各种媒体报道说，奥巴马的新闻发布会、演讲和非正式访谈都常用到“我”。一些有名的新闻分析师，如乔治·威尔、英国学者斯坦利·费希，甚至是临时的总统演讲稿撰写人，如佩吉·努南，都指出奥巴马不断使用“我”的情况。关于这个主题的一些文章发表在备受尊敬的地方——《华盛顿邮报》《纽约时报》，这里通常有勤奋的事实检查者。

唯一的问题是，没有人费力去统计奥巴马使用“我”的次数，或与其他人进行比较。从图10-2中可以看出，和其他任何一位现代总统相比，奥巴马使用“我”的频率最低。分析他的演讲也得到同样的模式。当奥巴马谈话时，他通常避免使用代词，特别是“我”。

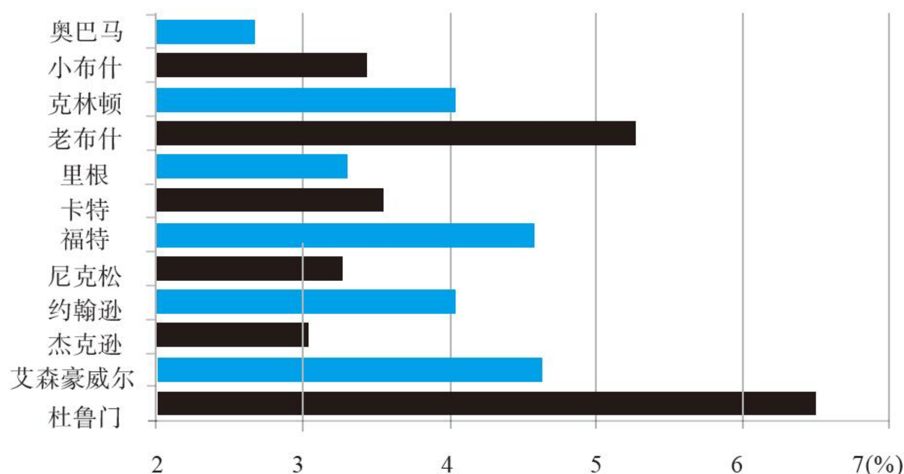


图10-2　新闻发布会中“我”字的使用

注：新闻发布会中使用的第一人称代词单数比例。

如果奥巴马比任何出现在脑海中的总统使用的“我”都少，为什么每个聪明的人会认为恰恰相反？问题可能在于我们自然处理信息的方式。正如我们在关于“我”的测试中发现的，大多数人认为那些自信的人使用“我”的频率比不安全或谦卑的人高得多。如果我们认为某人傲慢，大脑将搜寻证据来证实我们的想法。每当我们假定傲慢的人使用“我”字时，我们的大脑都注意到，啊，是寻找，额外证明他傲慢的证据。那些大肆宣传奥巴马过度使用“我”字的评论者，是不同意他政治观点的人，这不是偶然的。

还有另外一个故事。奥巴马使用“我”字频率很低，说明他是个自信的人。在2009年8月8日国家公共广播电台《周末版》（Weekend Edition）的访谈中，丹·巴兹和海涅斯·约翰逊被问及他们的书《2008年美国之战：一场非凡的选举》（The Battle for America 2008:The Story of an Extraordinary Election）。约翰逊回顾了自艾森豪威尔以来他的杰出政治报告生涯，他指出，奥巴马是“他所见过的最有自信的总统”。

奥巴马使用代词的方式证实了约翰逊的观点。与其他现代美国总统相比，自选举以来，奥巴马一直保持相对较少地使用“我”。与媒体专家的声明相反，奥巴马既不“过度喜欢”第一人称单数代词（如乔治·威尔斯所写），也不是在展示“帝王风范”（引自斯坦利·费希）。相反，奥巴马的语言显示出自信，同时显示出情感上的距离。

因攻击和战争避免使用“我”

从“我”字能看出人们关注的事情。如果一个人关注自我、缺乏安全感或谦逊，就会倾向于高频率地使用第一人称代词单数。如果一个人自信、专注于某种任务或说谎，他使用“我”字的频率下降。我们很少能够获得现实世界中几年内人们每天使用的语言，以便追踪他们关注的焦点和思维模式的波动。但是近代的美国总统（特别是布什）是例外。与大多数前辈不同，布什会见媒体的次数非常多。在他的第一个四年任期内，至少360个新闻发布会或媒体聚会被转录并发布在Whitehouse.gov网站上。通过分析他在回答问题时“我”字的使用，我们可以确定他对正在进行的政治和社会事件的想法如何。

布什的总统任期很可能是几代人的历史分析素材。第41任总统的儿子，作为一个官二代，他大概知道自己要成为一个热情和有魅力的人，但对于成为一位什么样的总统，他没有一个清晰的愿景。布什的总统之路比大多数人更加动荡不安。布什总统任职9个月之后，与乌萨马·本·拉登合作的飞机操作人员袭击了世界贸易中心和五角大楼，使3000多人遇难。不到一个月，2001年10月7日，布什指挥了一次攻击，推翻了阿富汗政府，但是未能捕获本·拉登。这是他作为总统期间最有争议的行为，布什将矛头转向伊拉克，声称该国藏有大规模毁灭性武器，并于2003年3月发动了全面入侵。从来没有发现任何大规模毁灭性武器，而美国及其盟国直到今天仍继续占领伊拉克。他的第二个任期也有问题，包括卡特里娜飓风摧毁新奥尔良、伊拉克的动荡以及全球经济崩溃的开始，这里不再讨论。

裹在布什行为外面的是他的性格。许多人认为他是相对透明的，有时表现出孩子气、暴脾气、防御性和同情心。在阅读他的新闻发布会记录时，经常出现他个性的不同方面，即使他是在回答同一个问题。无论你的政治派别为何，记录显示，他既温暖、迷人，偶尔又傲慢、刻薄。

三个相互关联的事件定义了他的第一个任期。第一个是“9·11”事件。正如上一章所述，在几天之内，美国人民从容忍他到崇拜他。第二个事件，世界没有直接见证，就是他决定入侵伊拉克。第三个是2003年3月的真正入侵。图10-3绘制了布什在每月的新闻发布会上“我”的使用情况。每个月，他通常和媒体对话6～8次，每次大约说1000个字。很明显，在他任职前三个月，“我”的使用有所下降。事实上，这种情况会发生在大多数新任总统身上（尽管奥巴马不这样）。开始担任总统工作通常是一个吓人的经历，需要几个星期去适应。

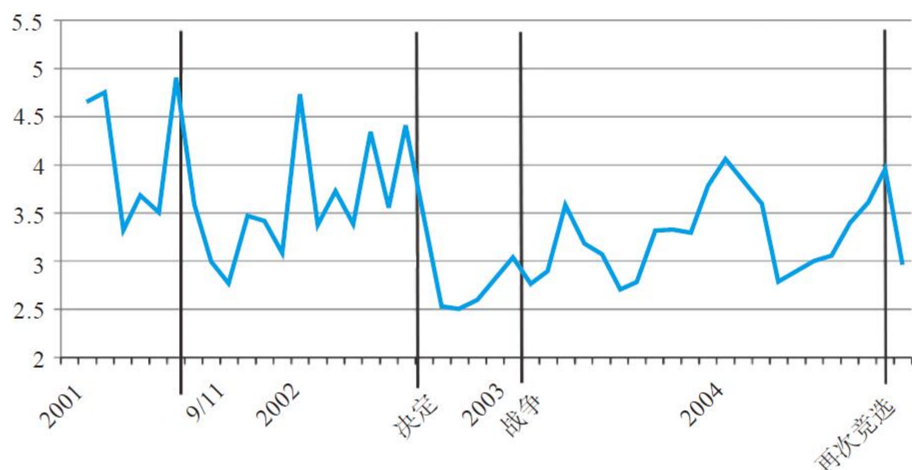


图10-3　2001～2004年，布什在新闻发布会上“我”的使用

注：乔治W.布什总统第一任职期间的360次新闻会议上使用的“我”的情况（基于占总词数的百分比）。纵轴表示以下内容：9/11＝9月11日的攻击事件；决定＝可能在伊拉克进行战争的最后决定（2002年10月）；战争＝入侵伊拉克（2003年3月）；再次竞选＝2004年11月。

“我”的使用第一次戏剧性的下降发生在“9·11”事件之后。和大多数美国人一样，布什减少了“我”的使用，增加了“我们”。攻击对他的影响是巨大的，在接下来的几个月里，他的注意力集中在一些紧迫的事情上，如入侵阿富汗、炭疽攻击事件，试图重新调整政府以处理恐怖主义。你会注意到，2001年11月以后，大多数月份他使用的“我”字逐渐增加。

“我”的第二次大幅减少发生在2002年9月中旬。据白宫学者说，布什政府一直受到伊拉克及其领导人萨达姆·侯赛因的困扰。此外，白宫内有些人认为萨达姆是“9·11”事件的幕后主使，并且他打算建立一个核武库。2002年6月，布什政府的一些人开始提出一种新的外交政策。布什原则（有些人这么称）的一部分，是美国有理由先发制人地攻击一个有敌意的国家。

整个2002年的夏天，一个预谋入侵的计划在秘密进行。9月底，布什政府要求国会授权发动伊拉克战争。这是一个讨价还价的工具，而且只有在伊拉克不允许检查人员全面进入该国搜寻大规模毁灭性武器时才使用的最后手段。国会于10月16日投票支持这项要求。在当天晚上的一场全国性电视演讲中，布什明确表示希望能够避免战争。马克·丹纳是《纽约客》杂志作者和加州大学伯克利分校的新闻教授，在他的一篇文章中，最终出现

英国文献证明，在美国国会的最后祝福下，发动战争实际是确定的事情。

想象一下，你是一个领导者，你知道你要攻击另一个国家。为了确保计划有效实施，你必须完全保守秘密。你不能透露任何关于部队行动、重新调整国家经济格局的信息，也不能让敌人知道你的计划是什么。你必须警惕你所说的话。不仅你的话是有欺骗性的，而且你必须关注政府的每一个方面，以协调秘密战争。要完成所有这一切，一个有影响力的领导者必须采取行动，而不是在那里考虑他自己的感受。

在准备发动威胁的人中，减少使用“我”字是一个强有力的信号。第二次世界大战期间，杜鲁门将原子弹投掷在日本国土之前，以及1939年希特勒入侵波兰之前，类似的模式出现在他们的语言中。

领导人的语言可以预测战争爆发的想法也曾被其他人提出。比利时心理学家罗伯特·霍格拉德研究了领导者在发言中如何谈及权力和联盟。当谈及联盟和友好关系很常见，而关于权力、侵略和征服的评论很少时，国家的前景通常是好的。但是，如果权力和侵略的话题增多，而与教育和关系相关的字词减少时，请小心。领导人之间的主题权力多/联盟少是战争的可靠预测。北爱尔兰、格鲁吉亚和俄罗斯的冲突以及中东的各种热点都是由各国领导人的语言变化推定的。

最后一个观察。有趣的是，战争开始前布什的“我”的使用几乎从来没有改变过。它短暂地上升了一点点，符合2003年5月他的“任务完成”演讲，当时他觉得战争结束了。随后2003年和2004年夏天“我”的使用下降，反映了他对战争的日益关注，2004年他对新闻界对其决定提出质疑具备防御性。

通过语言重看历史

我一直努力提出一个令人兴奋的可能性：语言分析可以运用到政治和历史的研究中。只要有语言痕迹（无论什么语言），计算机文本分析方法都可以帮助解释作者的心理。这方面的一些研究具有实际用处。其他研究纯属娱乐。

例如，为了更好地了解诸如基地组织等群体领导人的心理，我和我的学生、同事曾与几个联邦机构合作，试图了解他们的对外公开的信息和有时残酷的行动之间的关系。我们可以通过研究他们的语言或他们追随者的语言来理解他们的号召力吗？我们用这些方法来更好地了解在美国的极端组织，包括从极右翼的新纳粹到极左派的气象员派 [1] 。一般来说，我们发现暴力团体与非暴力团体有不同的语言指纹。此外，随着从非暴力到越来越暴力，他们的语言随之发生改变。

我们还将我们的语言工具用来研究过去和现在强大且通常是专制的领导人。例如，希特勒和其他领导人，当他们自己从革命者成为独裁者，成为有时被尊重的国家领导人，他们如何改变自己的语言使用？他们的语言如何预测他们国家的变化，以及事件如何改变他们的语言？

最重要的是，语言分析可以用新的方式阐明历史事件。我们从联邦党人文集、诗人和学者之间的关系，甚至披头士的例子中已经看到一些可能性。可以回答的历史问题仅受语言样本可用性和研究者想象力的限制。例如，澳大利亚探险家亨利·海利尔是真自杀还是可能被谋杀？（根据他日记中的文字，他可能是自杀）。《圣经》里圣保罗的文字真的是他写的吗？（不——根本不可能。）Lady Gaga真的和汤姆·克鲁斯有暧昧关系吗？（不知道。希望不是真的。）

现在轮到你了。

[1] 20世纪60年代美国的一个激进派别。——译者注

通过人们的语言推测他们未来的行为

语言能否告诉我们某个人是否会成为一个杰出的总统、合适的配偶、优秀的员工或学生？事实上，我们一直在脑海中想这些事情。如果一个学生写信给我，想在我的实验室工作，我会读她的电子邮件、她的简历和她的规划。她的语言反映了她的成就、对未来的期望以及她的性格，而这些将是我决定的基础。依靠在线约会网站的人最终可能根据他的语言使用而决定他们的配偶。是的，我们不仅考虑政治候选人的面部表情和肢体语言，而且评估他们演讲和用语言表达的计划。

虽然我们在投票、招聘或结婚之前会倾听和思考人们说的话，但是令人印象深刻的是，我们的判断经常会出错。语言分析程序能做得更好吗？或者它能帮助我们做出更好的决策吗？这还没有最终定论，但在计划上大学的学生和准备过正常生活的囚犯中会发现一些有趣的案例。

通过大学录取申请书的语言预测大学成绩

我和我的语言学家朋友大卫·比弗在一家酒吧讨论代词。（你听说过多少酒吧的故事是这样开始的？）如果语言使用和人们以后的行为有一些简单的关系，那不是很好吗？我们开始互相挑战，我们谁可以得到与重要的现实世界行为相关联的语言样本。然后我想起了加里·拉韦涅。

几年前，我遇到了加里，他是得克萨斯大学奥斯汀分校办公室的首席研究员。加里不是通常我们见到的统计学家。他已经出版了一系列关于大规模杀人犯的非虚构惊悚小说，以及最近一本关于得克萨斯州的学校废止种族歧视的历史书。此外，他还对什么因素能够预测谁会在大学取得成功感兴趣。得克萨斯大学总是拥有美国校园最大的学生团体。虽然学校每年招收超过7000名新生，但入学竞争非常激烈。申请过程的一部分涉及学生写两篇文章。

学生在入学申请文章中使用的功能词能否预测他们的大学成绩？这对于我和大卫来说都是一个具有吸引力的问题，事实证明，对加里也是。需要说明的是，这不是一个发明一种新方式来评估入学申请文章以确定谁应该被录取的战略。相反，我们首先想知道语言使用是否与学业成绩有关，如果有关，我们能否影响学生，让他们在大学成为更好的作家和思想家。

我们最终分析了来自2.5万名入学四年以上的学生的5万多篇文章。结果是明确的。语言使用确实与学生四年制的大学成绩相关。与良好成绩最相关的词语类别是：

·高比例的冠词和具体名词

·高比例的大词

·低比例的助动词和其他动词（尤其是现在时态）

·低比例的人称代词和非人称代词

这个词群应该看起来很熟悉。你可能还记得前面的章节中讲到，人们分类型思维和动态型思维的程度有所不同。分类型思考者倾向于关注物体、事件和类别。这个维度的另一端是更动态的人。动态思考时，人们在描述行动和变化。通常，分类型思考者将他们的精力投入到其他人身上（这解释了为什么他们高比例地使用人称代词）。

这是否意味着分类型思考者比动态型思考者更聪明？根本不是。然

而，美国的教育系统旨在测试人们对物体和事件分类的方式。

看看这两篇大学入学的文章，它们显示了分类型思维与动态型思维的不同。（这些文章的实际内容已大大改变，同时保持文章中名词、大词、动词和代词的使用比例。）

分类型思考者

选择的概念在西方哲学中发挥了重要作用。一个人的个性通过有意识和无意识的考虑被抛光到更加清晰的状态。如果没有对选择做出彻底控制，就不可能实现自由的最终目标。

父母离婚使我过上了双重生活。我有点逃避现实，这有严重的负面影响，包括不能理解其他观点……

注意作者如何有条理地定义和分类思想和经验。写作是结构化的，且很大程度上是个人化的，但同时又是沉重的。我们将分类型思考者与更动态的思考者进行比较。

动态型思考者

我看着哥哥，他比我大、比我聪明，我只是看着他哭。我还没意识到，我也在哭。我真的不知道为什么，但如果我的哥哥认为这是不好的事情，那就是不好的事情。所有人都在移动，但这是我的旅程规模，一个1130千米的旅行，从一个小农村到美国最大的城市之一。这将是一个挑战，但也是一个成长的机会。这对一个孩子来说，涉及放弃重要的一切，家庭、朋友、学校。

动态型思考者更加个人化，而且在讲述一个故事。语言非正式性更强，更简单，词更短。每个句子都有多个动词，这使得故事更加生动。

虽然这两个学生入学成绩都很好，并且都获得了文科学位，但大学期间，分类型思考者的成绩每年都比动态型思考者高很多。这不是因为分类型思考者是一个更好的作家。相反，分类型思维方式更符合我们在大学所获得的东西。例如，大多数考试要求学生将复杂问题分解。同时，很少有课程要求学生讨论正在进行的事件或讲述自己的故事。

大多数大学不会使用字数统计程序来决定录取谁。一旦学生发现学校在使用这样的系统，那么他们的申请文章里会出现大量的大词、冠词，并且几乎不用动词。相反，这些发现指向我们可能考虑的培训方式，用来培训高中生和更小的学生。就我们的教育系统鼓励和奖励分类型思维的程度而言，应该明确培训学生这样。

还有一个论点是，我们应该探讨是否要在大学层面鼓励动态型思维。讲述故事和追踪人们生活中变化是对人们来说很有用的技能。同时出现另一个问题，大学成绩好的人毕业后情况如何。是否有可能动态型思考者能更好地适应或更快乐呢？以及，思维方式有多灵活？我们所有人时而需要分类地思考，时而需要动态地思考，这是完全可能的。

通过语言预测出狱后的生活

毒品滥用和酒精滥用对社会造成巨大损失。许多国家试图遏制和治疗滥用的一种方式是建立治疗性社区，这些治疗性社区基本上是治疗性监狱，人们在被判定犯有毒品相关罪行后，有机会接受为期几个月的密集的戒毒康复和酒精康复。如果参与者成功完成该计划，并在释放后的指定时间内保持不沾毒品或酒精，那么他们的记录通常会被清除。大多数治疗性社区需要集中的团体治疗，并配合写作练习。

我之前的研究生安妮·瓦诺进行了一个宏大的项目，目的是了解女性在治疗机构内的写作方式是否可以预测她们从监狱释放后的生活。安妮和一个治疗社区合作，收集和转录了大约120名女性的写作样品。她关注的是女性在被释放后一个星期内写的文章。她预测文章是与个人相关的并且是真诚的。随后的几个月里，安妮与监狱长办公室合作收集后续信息，例如女性们保持工作的能力，以及她们是否在假释期间违纪或被逮捕。

不论以什么标准来衡量，女性们写的故事都非常有感染力。她们经常描述作为身体伤害和性虐待的受害者的例子，同时，她们详细描述自己对待他人的糟糕行为，例如她们的孩子。她们经常表达对自己离开监狱回到不确定的家庭生活的担忧。

离开监狱之后，120名女性中有15%被捕，另外10%在该项目完成后四个月在假释期逃跑。大约65%的人保住了稳定的工作。

有趣的是，女性最后一篇文章的写作方式预测了四个月后她们是否正常生活。与治疗成功密切相关的两个语言维度是：

·高社会-情感风格，包括使用人称代词和情绪词汇

·高比例的积极情绪词汇

离开治疗社区的女性的任务是融入新工作，融入一个运转的社交圈子。分类型和动态型思维对于这些女性来说只是不相关的维度。要在监狱外的世界中生存，她们需要关注别人和自己。看上去，她们在文章中展现的社会-情感风格和乐观风格是她们能够在外面用到的技能。

***

很难再想象出两个研究项目，可以像大学入学申请书和治疗性社区如此不同。对于某个团体，分类型思维预测更高的大学成绩；对于另一个团

体，社会-情感型语言预测更低的再被捕概率。语言的不同方面连接着我们生活的不同部分。

我喜欢这两个研究（实际上，我喜欢本书中的所有研究）是因为，隐性词在不同配置中重新排列，预测了一系列广泛的行为。例如，使用高社交-情感风格相关的语言可以帮助你待在监狱外面，为你当选总统做出贡献，还可能帮助你获得创作畅销歌曲的技能。

一方面，根据上下文，使用“我”字频率较高可能意味着安全感缺失、诚实和抑郁倾向，也可能意味着你不打算在近期发动战争。另一方面，使用“我”字频率较低可能会被大学录取，提高你的平均成绩，但可能会让你失去结交亲密朋友的机会。

回到已经出现过几次的话题很重要。与社会和心理状态相关的词汇反映了状态，而不是起因。它们反映人们脑海中发生的事情。她们从监狱释放之前使用高比例的人称代词和情绪词汇，她们在以社会-情感的方式写作。她们投入的治疗项目是否会真的让她们以社会-情感的方式思考，这是未知的。我们同样不可能知道，写作样本中使用的词是否直接影响她们被释放后的行为。更不可能的是，如果她们强迫自己在文章中使用这些词汇（认为这可能对她们有好处），是否会影响她们的狱后生活。

我们正站在一个新世界的门口。想想计算机对功能词的分析已经开启的许多应用。通过分析就职演讲或前辈的日记，我们就能够了解过去有影响力的作家或演讲者。我们也可以开始回答日常生活中一些重要的心理学疑问。我们可以了解我们的在线约会对象如何看待我们，区分哪些说唱艺术家是真正的黑帮，判断我们的治疗师是否像我们一样抑郁，或揭发那些认为自己享有最高地位的同事。

功能词可以稍微帮助我们更好地了解我们的世界。功能词的作者识别可以帮助抓捕罪犯或确定历史作者，功能词可以帮助了解总统或暴君的思想，还可以帮助预测人们未来的行为，功能词是人类心理的线索。然而，最有望成功的是，通过观察我们自己的功能词，可以开始更好地了解我们自己。

附录　识别和解释功能词的便捷指南

不同的功能词反映了不同的心理状态。代词、冠词、介词、动词和情感词汇为研究人们的思考方式、感受和社会关系提供了不同的信息。在未来所有有人使用语言的场景中，你都可以随意参考下面语言指南的任何部分。

重点关注：代词、动词时态

语言体现着我们的关注点。我们会谈论或记述心中所想的物体、事件和人。有一条好的经验法则是，非常关注他人的人倾向于使用很多人称代词，迷恋过去的人使用过去时态的动词。反过来看这一观察，通过计数人们使用代词的情况和使用动词的时态，可以猜测自然状态下人们在关注什么。

◎人称代词

我们花费在想念和讨论他人上的时间应该超过了其他任何事情。如果一个人令我们感到非常开心、极度愤怒或者极度悲伤，我们通常会不停地想他。我们会在谈话中不经意地和别人提到他的名字，当提及这个人时会甩出很多代词。因此，如果说话者在想念或谈论一个朋友，预期他会使用高频率的第三人称代词单数。如果说话者担忧右翼电台主持人或者官僚主义者，他使用“他们”的频率会高于平均数。

“我”字也是一样。如果人们处在不自在的状态中，他们的关注点会时不时地转到自己身上，虽然很短暂，频率却比自在状态的人高。比如说，人们在镜子面前做一份问卷时，使用“我”的频率比不在镜子面前要高。如果人们因为生病、痛苦或极度低落而更多地将关注点放在自己身上时，也会使用更多的“我”。与之相反，当一个人沉浸在一项任务中时，使用“我”的频率会比较低。

◎动词时态

人称代词体现人们的关注主题，而动词时态可以体现人们是如何思考时间的。不出意外地，当人们考虑过去时，使用过去的时态；当人们考虑未来时，倾向使用将来时。更有趣的是人们在这两种时态间转换的时候。例如，患有创伤后应激障碍的人在经历了可怕事件几个月或几年后常会情景重现。一个越战老兵这样描述发生在几十年前的可怕夜晚：

我们用推土机建造……沙坑……防止遭到零星炮击。我无法在沙坑中入睡（它像地穴）。今晚……我们连受迫击炮的攻击，突然我们的副排长崩溃了，蜷缩在沙坑的角落里歇斯底里地哭了起来。炮轰停下的间歇，我走到外面的护堤顶部。呸。外面一片漆黑。我们坐在那里，时不时受到零星的地面火力攻击，然后我们也会进行回击。……几个回合下来，也许一个小时以后，我站在护堤顶部瞭望，然后再次感觉到了他的存在。我盯着他看，尝试看清他的动作，但是外面太黑了。我站起来，跑去……要照明灯。我回到顶部，继续盯着，等着光亮，没有人来。我知道他在那里，而且不是一个人……显然，我很好。我只是不记得后来发生了什么。

请注意讲述者是如何在过去时和现在时态间转换的。副排长、炮轰、炮火声都发生在过去。“也许一个小时以后，我站在护堤顶部……”老兵回到了越战的现场，仿佛身处其中。他的动词时态告诉我们他的思维如何运转——越战经历依然持续成为他生命经历的一部分。

社会关系和情感：代词、冠词、情感词汇

语言最基本的功能是沟通。功能词为了解社会进程提供信息，社会进程对于了解他人来说非常必要。谁更有地位，一个团体是否工作和谐，某个人是否正在欺骗他人以及一段关系的亲密程度如何，这样的问题，功能词都可以提供一些线索。

◎地位和权力

在几乎所有灵长类动物的交往中，首要任务就是建立统治和地位。当然，我们人类完成这个任务的方式更为文雅。通常，我们不会在遇见他人时威胁他们。相反，我们会改变我们的姿势、语音语调和使用文字的方式。正如第7章所讨论的，人称代词的使用可以很快显示出地位，尤其是第一人称和第二人称代词，如“我”“我们”和“你”。与大部分人认为的正好相反，地位高的人比地位较低的人使用“我们”和“你们”的频率更高。而且低地位的人过度使用“我”。

◎诚实和欺骗

此外，我们说实话和说谎话的时候，使用功能词的方式不同。在说实话的时候，我们倾向于“拥有”所讲内容。也就是说，讲真话的人更容易使用“我”和“我的”这样的词。在说谎的时候，我们会疏远所讲内容。比尔·克林顿总统声称自己没有“和那个女性”发生性关系的例子是个令人吃惊的例子。“那个”女性和“我的女性”或者简单的“莫妮卡”相比，距离肯定更远。为人真诚时，我们讲话和思考的方式会更复杂。正如第6章所看到

的，撒谎不是一件容易的事情。如果不得不虚构一个故事，我们会避免使用某些连词（如“但是”和“或者”）、介词（“除了……以外”“没有……”）和否定词（“不”“从未”）。

◎陌生人和朋友

认识很久的好朋友之间使用语言的方式和两个陌生人之间非常不同。有些语言标志很明显：和陌生人之间相比，彼此相互喜欢的两个人使用更多的第一人称复数（“我们”“我们的”），还使用更多的积极情绪和消极情绪词汇。有共同背景的人之间更倾向于使用具体冠词“这个”，例如“这个椅子”，因为很有可能这两个人已经讨论这把椅子很多年了。陌生人之间最初会说“一把椅子”，尤其是在认识之初。

思考方式：连词、介词、名词、动词和因果词

如果你发现你最好朋友的配偶有外遇，你会告诉你的朋友吗？为什么告诉或者为什么不告诉？

诸如此类的问题迫使人们去思考复杂问题，这些问题往往没有简单的答案。人的答案通常需要一定程度的逻辑思维、推理性和因果思维。

意料之外的复杂问题还需要解决一个难题。人们通常以一种方式回答问题，然后从另外一个角度去评估这个想法是否合理。当人们在记录复杂问题时，他们使用的功能词为他们通常情况下的思考方式提供了线索。

虽然分析思维方式的方法有几十种，但是有三种特别适合于语言分析。

◎复杂思考VS简单思考

那么，你发现你最好朋友的配偶有外遇了。假设问12个人他们会怎么做，你会得到12种不同的回答。然而，有些答案体现出更为复杂的思考方式。

复杂思考者： 首先，他们的感情关系如何？如果这对伴侣对婚外关系已经有协议，那么没什么需要说的。告诉朋友还是保留秘密，要考虑到朋友或者其配偶所要付出的代价。如果反过来，我的配偶有外遇了，我会想要朋友告诉我吗？

如果朋友的配偶故意要伤害我的朋友，那我很有可能什么也不会说。但是我的确需要更多的信息才能回答这个问题。

简单思考者：发生在拉斯维加斯的事情就让它留在拉斯维加斯。这是我的格言。会发生的早晚会发生。我不会干涉。眼不见，心不烦。懂我意思吧？

尽管复杂思考者和简单思考者得出的结论一样，但是复杂思考者从不同的角度权衡了不同的选择。另外，复杂思考者依靠理智、逻辑，甚至是情感意识。注意他们的语言区别。复杂思考者通常会用大词、长句和更复杂的句式，常使用介词（“和……”“……的”“到……”）。顺便说一下，介词是很好的语言标记。它们帮助我们形成时空概念。

更重要的是，复杂思考者会区分时间和空间。“朋友的配偶曾经有外遇，但是已经都过去了”比“他的配偶有过外遇”更复杂。为了进行区分，说话者必须告诉我们哪些属于同一类别，而哪些不属于这一类别。“它属于这个类别但不属于那个类别。”为了做区分，有必要利用一组特定的词汇，我们称之为专属词。专属词包含“除了”“除……以外”“没有……”“除非”，以及一些相关的介词和连词。

不要忘了还有“如果那么”短语，这是逻辑思维的中心。如果一个人使用“如果”，那么根据定义，这个人正在做区分。

◎因果思考VS非因果思考

如果你醒来发现一条腿被截肢，会怎么想？我问过上百个大学生这个问题，答案非常吸引人。有些学生说，那是毁灭性的打击，会考虑自杀。有些耸耸肩说，那不会对他们的生活带来什么太大的改变。然而还有些学生关注的是，为什么他们会行走在汽车中，又是什么原因会使其中一辆车后退。

所有人都会很自然地进行因果思考。但是，有些人会更常进行因果思考。比如说，有些人会执着于为什么坏事会发生在他们身上。为什么我要在那些车中间走？为什么我不得不失去我的腿？在一系列令人惊叹的研究中，心理学家卡米尔·沃特曼和罗克珊·西尔韦采访了上千人，问他们如何应对生活中大的起伏，如乱伦、孩子夭折、丧偶。那些对可怕事件的因果性解释比较简单的人往往应付得很好。还有些人连因果性解释都不去寻找，他们也同样应付得很好。而另外一些人是最有困难的，他们总是拼命寻求答案，想弄清楚为什么会发生这样的事情。他们总是问：“为什么会这样？为什么是我？”但是从来不找答案。

大多数情况下，因果思维是非常宝贵的。如果你能为一件事找到令人满意的原因，今后发生类似事件，你就可以更好地应对。但是，如果你无法找到问题的答案，继续寻找可能只会给你带来挫折和不快。

捕捉语言中的因果思考比较容易。回忆一下高中英语知识，连词是用来连接词和短语的，例如“或者”“而且”“但是”。有些连词专门用来表示因果思考，例如“因为”“因此”“所以”“既然”。有大量的名词和动词，虽然它们不是功能词，但是也表示因果关系，例如“引起”“影响”“原因”“原理”“驱使”“抑制”。自然状态下，人们写作或说话时，越多地使用这些词汇，就越会搜索或考虑因果。

◎动态型思考VS分类型思考

在第2章，人们被要求描述图中的两个人。其中两个描述如下：

学生1：在上述画面中，一个老年妇女正要和一个看上去高人一等并精于算计的中年女子说话。

学生2：我看到一个老妇人在回想她当年年轻美丽时候的样子。

第一个学生，男生，在描述这个女性时使用的词汇直接且具体；与之对比，第二个学生，女生，描绘了一幅动态画面，要开始构建一个故事。在看待物体、事件和特殊人群方面，男性看待世界的方式更具有分类性。分类思考的方式需要使用特定的或具体的名词。伴随特定名词一起出现的是冠词。

在日常交流中也可以看到动态型思考和分类型思考的区别。约翰·麦凯恩和巴拉克·奥巴马在2008年秋天的总统竞选中，显示出截然不同的思考方式。一方面，麦凯恩在辩论、采访，甚至是演讲中，使用冠词的频率非常高。而另一方面，奥巴马使用冠词的频率比麦凯恩和其他任何有竞争力的总统候选人使用冠词的频率都低。10月中旬，这两个男性进行了他们之间最后的辩论。两人都被要求解释为什么美国的教育体系是世界上最昂贵的，却没有被看作工业化国家中非常强大的。

麦凯恩：嗯，这是21世纪的公民权利问题。毫无疑问，经历了漫长而又艰难的努力之后，在美国，我们已经取得了平等入学机会。但是在一个低收入区域，把孩子送去一所不合格学校是你唯一的选择，这优势是什么呢？所以，学校间的选择和竞争是一个关键因素，这在新奥尔良和纽约已经得到证实。

奥巴马：这可能更与我们的经济前景相关联，意味着它有国家安全影响，因为地球上没有一个国家在经济衰退的同时还继续保持其作为军事大国的首要地位。所以我们有必要使我们的教育体系更加完善。现在，关于更多钱还是更多改革的问题存在争论，我认为我们两者都需要。某些情况下，我们不得不进行投资。

麦凯恩在回答时，将问题分解成了一个个部分，使用的冠词数量几乎是奥巴马的两倍。一方面，在麦凯恩看来，教育是一个公民权利问题，是奋斗的一部分，无法通过把低收入家庭的孩子送入不合格学校来解决。可能会有各种解决方法，但这不是其中的一种。从另一方面讲，奥巴马回答方式更灵活——与变化的过去和未来相连。麦凯恩考虑问题比较具体，而奥巴马考虑问题更抽象，依据的原则更广泛和多变，如经济前景、教育体系、货币、改革和投资。

回过头来看这些思考方式（复杂VS简单，因果VS非因果，动态VS分类），没有哪个思考方式比另外一个更好或者更有效。有时候复杂的、因果的和动态的思考方式可以帮助人们度过这一天；有时候这些思考方式可能会成为麻烦。事实上，我们所有人都在各种思考方式间往复，这取决于我们在思考什么。奥巴马在回答关于教育的问题时有些抽象，但是并不能因此判断出他在吸烟方面的决定或者他是否决定吃阿司匹林来缓解头痛。

注释

辛迪·钟、山姆·高斯林和露丝·彭尼贝克为本书提供了大量评论和反馈，除了他们，还有一些人读过部分或全部原稿，包括大卫·毕佛、莫莉·艾尔兰、杰夫·汉考克、莫琳·奥沙利文、里奇·斯拉切和依拉·陶兹斯克。感谢所有人。

前言

1.一般人的词汇量大约在3万~10万之间不等。一部分问题在于对词的定义。例如，包括单复数吗？同一个词的不同写法呢？或者你使用的词可能字典里没有，比如脑补（意思是，在脑内幻想希望而没有发生的情节）。关于功能词数量的讨论，参见钟和彭尼贝克（2007）的文章。

2.语言分析中最令人兴奋的突破之一就是最近发布的谷歌图书语料库。让·巴蒂斯特·米歇和同事们合著了一篇令人惊叹的文章，被刊登在《科学》杂志中，文章分析了500多万本书中的文字，也就是所有已经印刷的书籍的4%。作者以过去200年的语言使用为研究重点，发现很多历史趋势。例如，弗洛伊德或达尔文被提到的频率是什么？名声能持续多久？语言是如何发展的？

3.一些结合了基本社会问题和心理问题的关于语言的畅销书包括：欧文·戈夫曼的《谈话形式》（Forms of Talk）、乔治·莱考夫和马克·约翰逊的《我们赖以生存的隐喻》（Metaphors We Live By）、乔治·米勒的《语言科学》（The Science of Words）、史蒂芬·平克的《语言本能》（The Language Instinct）、德博拉·坦嫩的《你误会了我》（Your Just Don't Understand），以及安娜·维兹毕卡的《通过关键词来理解文化》（Understanding Culture Through Their Key Words）。

第1章　最容易被人遗忘的词汇的秘密

1.Analyzewords.com网站是我们开发的几个实验网站之一。它是三方协作的结果：Slate.com网站的作家克里斯·威尔逊；我的女儿蒂尔·彭尼贝克，她是世界级社交媒体专家；长久以来我的电脑大师罗杰·布斯。该程序可以分析任何人近期发布的推特文。通过几年来我们开发的用来研究语言和性格的算法，我们可以通过语言使用来估计一个人的性格。与所有实验性系统一样，请不要过于重视结果。

2.表达性写作研究有着丰富的历史。基本思想是，如果人们被要求连续三四天每天记述创伤性经历15~20分钟，那么与那些被要求记述表面话

题的人相比，他们的身体健康和精神健康后续会有所改善。这种方法的效果一般但是效果稳定。为了更好地理解这本书，请参阅我的畅销书《打开心扉：情感表达的治愈功效》（Opening Up），以及泛达罗利（2006）、彭尼贝克和钟（2011）合著的文章。

3.如果你想要了解LIWC程序开发的技术信息，请阅读这篇与“语言获得和词汇计数”相关的文章（www.psy.utexas.edu/pennebaker ）。LIWC程序可以在下面的网站通过商业方式获取：www.liwc.net 。销售该程序获得的所有利润都回馈给得克萨斯大学奥斯汀分校心理学系的研究生课程。

LIWC绝不是唯一的文本分析程序。现在有几十种不同的程序可用。对语言和心理学感兴趣的人来说特别有用的程序包括：

阿特·格雷泽的Coh-Metrix（cohmetrix.memphis.edu），它可以计算任何文本组的可读性和一致性。

罗德·哈特的Diction（www.dictionsoftware.com ），旨在掌握信息的口头语气，尤其关注政治演讲。

汤姆·兰道尔的Latent Semantic Analysis（lsa.colorado.edu），这套程序允许人们比较任何文本之间的相似性。

迈克·斯科特的Wordsmith（www.lexical.net ），一款非常棒的通用词汇分析工具。虽然不是专门为高级心理分析而设计，但它是一个很好的词计数软件。

4.关于积极情绪、消极情绪和认知词汇变化的发现首次发表于1997年，作者是彭尼贝克、梅恩和弗朗西斯。另见摩尔和布洛迪（2009），格拉哈姆、格拉泽、洛文、马拉基、斯托厄尔和凯寇尔特-格拉泽（2009）的文章。

5.代词和变化观点的发现基于我和夏洛克·坎贝尔于2003年出版的一篇文章。

6.永硕·金的一篇重要论文记录了关于表达性写作有趣的发现之一。在她的表达性写作项目中，韩语-英语和西班牙语-英语的双语学生有的用他们的母语，有的用第二语言，有的两者都用。永硕让所有参与者在写作的前两天和写作实验后一个月携带便携式磁带录音机。与没有写作的人相比，记述情绪动荡的参与者花费更多的时间与他人相处、交谈，笑声也更多。虽然效果不大，但是既使用母语又使用第二外语写作的人往往受益最

大。

7.切瑞·休斯和玛莎·弗朗西斯是南卫理公会大学最早获得心理学博士学位的两名学生。尝试通过影响人的词汇选择来改变他们想法的实验结果是休斯1994年博士论文的基础。最近，我和易泰熙、辛迪·钟要求人们在表达性写作中以第一人称（好像他们只是在描述自己的经历）、第二人称（好像他们在对镜子中的自己叙述）或第三人称（好像他们正在看电影中的自己一样）写作。在另一项研究中，要求人们在不同的文章中使用不同的视角。令我们惊讶的是，参与者表示，从任何角度写作都是有帮助的，转换角度的写作也是有帮助的。我们目前的想法是，强制视角转换对于一些人来说可能是有益的，但对于其他人来说不一定。这不是我们正在寻找的那种突破性的结论。

第2章　重要的不是内容，而是语言风格

1.本书中，有些引用来自我的学生或实验对象，有些来自互联网上的文字，甚至有些来自朋友或家人的对话或电子邮件。所有引用中具有识别性的信息都已被删除或更改。

2.本书中，“风格”“功能词”和“隐性词”交替使用。它们还有许多其他的名字：垃圾词汇、小品词和封闭类词。语言学家倾向于区分这些相互重叠术语的精确定义。

3.功能词使用频率表基于我们自己的数据分析，语言样本来自2000多个对话、200篇小说、2万个博客以及数千个大学生和成人写作样本。想要了解更多信息，请参阅彭尼贝克、钟、艾尔兰、冈萨雷斯和布思（2007）的文章。

4.政治运动是非常好的分析对象。总统候选人每天都是最受关注的焦点，无论他们在演讲、做访谈，还是只是在和街上的人交谈。他们的大部分话语都被转录，使文本分析专家能够轻松追踪政治运动过程中的语言。对于2004年美国总统大选的一些分析，参见我和里奇·斯拉切、辛迪·钟、洛里·斯通合著的文章。

5.讨论大脑和语言角色的最好的书之一是乔治·米勒的《语言科学》。这本书对于那些对该领域没有大量研究的人来说，是对该话题最完美的介绍。

6.乔治·米勒的书中讨论了菲尼亚斯·盖奇的故事。亚历山大·卢里亚在他1973年的经典书中描述了巴甫洛夫与他的狗的故事。

7.在既定对话中，基于以前的共同知识，双方能够知道代词和其他词汇指的是什么，就是所谓的共同点。它对于交流非常重要，所以各个领域都对其进行了广泛的研究。心理学界，赫布·克拉克作为共同点理论来研究；语言学界，大卫·比弗作为预设理论来研究；认知科学界，菲尔·麦克卡锡和同事们作为已知性/新奇性来研究。

8.不幸的是，跨语言研究在本书中没有详细介绍。我的前学生奈拉·拉米雷斯·埃斯帕萨发现，人们从英语切换到西班牙语时思维方式发生了一些显著变化，反之亦然。例如，双语者觉得他们在讲英语时更外向，而在讲西班牙语时更拘谨，尽管他们的行为表现相反。与辛迪·钟等人合作，奈拉发现人们用西班牙语和用英语思考同一个话题时，想法会略微有些不同。另见鹿岛（1998）。

我可以肯定，任何人写有关语言的文章都会提到“沃尔夫假设”（如果你是一位人类学家，那就是“萨丕尔-沃尔夫假说”）。沃尔夫假说的极端版本是：一个人的语言和词汇量决定了这个人的认知。所以，比如说，如果我不能用语言描述橙色，那么我就无法感知橙色。自20世纪60年代开始，由于心理学和语言学的一个有趣的历史运动，沃尔夫假说被彻底摒弃。新的和更有趣的版本不断出现。奈拉、莱拉·博罗帝茨基等人的研究表明，在某些情况下，我们的语言可以引导我们的注意力、思维和记忆。

第3章　语言体现性别、年龄和权力

1.我们对性别差异和语言研究最好的总结在我和马特·纽曼、卡拉·格鲁曼、洛里·海德曼2008年发表的一篇论文里。人们坐在镜子前完成一份问卷，与没有镜子时相比，他们使用更多的第一人称代词单数（Davis and Brock，1975）。在芭芭拉·弗里德里克森、汤姆·安·罗伯茨和他们的同事们的一系列研究中，证明了女性自我关注的本质。温蒂·加德纳和她的同事们展示了人们对代词工作方式的刻板印象。

2.问任何文化背景中的任何群体，他们都会认为女性说的话远比男性多。我和我以前的学生马蒂亚斯·梅尔、塞名·维齐尔、奈拉·拉米雷斯、里奇·斯拉切在美国和墨西哥进行了六次实验，我们邀请了将近400名大学生穿戴数字录音机2～4天。所有他们说的话后期都被转录为文字。令人惊讶的是，所有实验中，男性和女性每天讲的词数几乎完全相同。一项没有可发表结果的研究是非常罕见的，但2007年《科学》杂志还是发表了它。

3.博客项目于2007年和2009年在我和什洛莫·阿葛曼、莫西·科佩尔、乔纳森·施勒合著的论文中发表。

4.我和莫莉·艾尔兰的论文目前正在审查中。

5.除了彭尼贝克、格鲁曼、勒夫和达布斯的文章外，另见乔·马纳的实验室以及罗伯特·约瑟夫斯及其同事关于睾酮和行为的近期研究。

6.随着年龄的增长，我们的个性也会发生变化。理查德·罗宾斯和他的同事们展示了更高的自尊心，约翰·莱欣和尼克·马丁研究显示情绪稳定性有所增加，赛德尔等人（2009）则发现，随着年龄的增长，关系代词使用增多。本节的许多内容都是基于彭尼贝克和斯通发表于2003年的文章。

7.过去的20年里，关于社会阶层和身体健康方面的研究在美国大幅度增加。一些关于这个话题的优秀文章包括艾迪斯·陈、凯伦·马修斯和托马斯·博伊斯，以及南希·阿德勒和同事们的经典作品。

8.几乎没有完整的科学研究追踪不同社会阶层的人在家中的自然语言使用。哈特和里斯利的研究很有希望，正如布兰迪斯和亨德森在1970年的研究一样。

第4章　性格：发现人的内在

1.大多数关于个性和语言的初步工作是与劳拉·金合作的结果。你可以通过阅读彭尼贝克和金合著发表于1999年的论文来了解实际分析。一些人扩展了这项研究，特别是语言与自我认知的个性如何关联的问题。两个伟大的项目已经出版，一个由弗朗索瓦·迈雷斯及其同事完成，另外一个由塔尔·雅克尼完成。心理学家丽莎·法斯特和大卫·冯德，以及我和马蒂亚斯·梅尔、山姆·高斯林发表了关于词汇使用和性格的文章。另请参见乔恩·奥博兰德、阿拉斯泰尔·吉尔和斯科特·纽森的创作。

2.辛迪·钟对意义提取方法（MEM）的见解已被世界各地的实验室采用。基本上，研究人员可以使用它自动从文本中提取主题。最有用的应用之一是从不同语言编写的文本中提取主题。例如，我们可以让计算机分析上千个博客，比如芬兰语的博客。使用MEM，我们可以提取在芬兰博客中受欢迎的几个主题。所有我们需要做的只是在提取结束后，找一个说芬兰语的人，让他来告诉我们，我们发现的主题是什么。第一篇意义提取论文由钟和彭尼贝克在2008年发表。

3.对主题统觉测验的一些评论反映了对于罗夏的评论。例如，传统基于评委的评作文方法是主观的。两个人看同一篇文章，可能会有完全不同的理解。还有一个问题是，当你要求同一个人两次描述同一张图片时，他们通常会想出一个新的故事，通常会是不同的主题。传统的性格研究人员想要得到不同场合相同的测试结果。不幸的是，人们太有创意了——他们喜欢创造新的故事，特别是当他们需要告诉同一个人时。

4.本·拉登项目由彭尼贝克和钟于2008年发表。我和辛迪·钟、阿特·格雷泽、杰夫·汉考克、大卫·比弗都在沉迷于一项更广阔的研究，以了解领导者如何随着时间推移改变他们的言辞。希特勒以及不太极端的西方领导人在战前、遭受袭击之后和政权开始败落期间，都存在语言使用的变化。

第5章　情绪测定

1.许多人认为情绪会影响我们的思维方式。例如，芭芭拉·弗雷德里克森提出了一个积极情绪构建模型。她的研究表明，当人们高兴时，他们能够回到常态，看到更广阔的视野，从而形成新的想法。托马斯·博尔卡维奇发现，悲伤或沮丧的人倾向于将消极的情绪集中在过去的事件上，而焦虑的人更担心未来要发生的事件。最近，查尔斯·卡佛和埃迪·哈蒙·琼斯提供了令人信服的证据证明，愤怒的情绪能够激活与近似积极情绪相关的大脑区域。悲伤或焦虑的感觉通常与逃避相关的大脑区域相关联。

2.过去30年来，几个实验室独立发现了抑郁与自我关注之间的联系。沃尔特·温特劳布是词分析创始人之一，他是第一个报道语言行为的人，请参见他于1981年发表的书。汤姆·塞兹斯克和杰夫·格林伯格发展了一个自我关注引起抑郁症的理论。除了我和香农·斯特尔曼于2001年发表的自杀诗人项目之外，鲁德、葛特纳和彭尼贝克2004年发表的一篇论文发现，与不抑郁的大学生相比，抑郁的大学生在描述大学生活的文章中使用更多的第一人称代词单数。另见考夫曼和塞克斯顿最近关于自杀诗人的自我关注的著作。

3.朱利亚尼项目的参考文献是彭尼贝克和雷2002年的文章。

4.某种程度上，丹·韦格纳对思想抑制的研究非常巧妙。根据托尔斯泰童年的经历，韦格纳要求学生在短短一两分钟内不要想一只白熊。这么做，他发现实际上他们开始频繁地想到熊。更令人震惊的是，当学生被告知他们不再需要压制他们的白熊想法时，许多学生说想到白熊的频率更高。韦格纳的早期研究使他开发了一系列有关思维监控信息的重要理论。

5.一些调查人员使用博客和其他社交媒体来追踪大型社区人们的情绪状态。例如，彼得·多德斯和克里斯托弗·丹福斯用它追踪幸福，还有亚当·克拉默；心理学互联网研究先驱鲍勃·科罗特，研究在线社区的公告留言板来追踪情绪支持；我和伊丽莎白·里昂、马蒂亚斯·梅尔分析了支持减食欲药物者的博客来评估他们的情绪概况；马库斯·沃尔夫、辛迪·钟和汉斯·科尔迪分析了心理治疗患者发给他们治疗师的电子邮件。

6.几年来，我和我的学生们研究了几个大规模动荡。这项工作大部分是试图了解大规模的情感事件随时间的推移如何展开。要了解更多关于它

们的信息，请参阅以下相关论文：得克萨斯农工大学篝火灾难（Gortner and Pennebaker，2003）；戴安娜王妃的死亡（Stone and Pennebaker，2002）；洛马普列塔地震和波斯湾战争（Pennebaker and Harber，1993）；2001年“9·11”事件（Cohn，Mehl，and Pennebaker，2004；Mehl and Pennebaker，2003）。

7.圣海伦火山爆发后，我和同事达伦·纽森和在俄勒冈和华盛顿附近游走了约一个月。在经历了大规模损伤的城镇中，我们随机采访人们和打电话给他们。一个有趣的发现是，社区受到的损害越多，居民越想要谈论。那些靠近火山且逃脱伤害的社区，对外界最具防备心，最不愿接受访问。请参阅我之前的书《打开心扉：情感表达的治愈功效》，以了解更多详细信息。

8.社会心理学中特别有前景的研究有关“情感预测”。据其创始人丹·吉尔伯特、蒂姆·威尔逊等人介绍，人们预测他们对未来事件做出何种反应的能力很差。一般来说，人们想象他们受到创伤后的抑郁情绪会很严重，而实际遇到后通常没有想象中那么糟。

9.辛迪·钟的论文涉及186个博客使用者一年内在博客社区讨论节食。

10.秘密总是有害的吗？其实不是。事实上，秘密可能有害，但将秘密告诉他人有时会更糟。关于秘密的心理学精彩总结，请参阅安妮塔·凯利的书和文章。

11.劳拉的故事在我的书《打开心扉：情感表达的治愈功效》中首次被描述。

12.社会经常采取出于好意却无效的干预措施。通常，除了集体晤谈（CISD），禁欲型性教育和大胆教育引发的问题比解决的问题还多。见蒂莫西·威尔逊的《转向》（Redirect）一书。

13.人们有强烈的冲动来谈论情感事件。伯纳德·里姆和他的同事们发现，所有文化中的人们通常都会向他人吐露几乎所有类型的情感事件。我最喜欢的发现之一是，如果你有一件感情方面的事告诉了朋友，然后让朋友保密，那么你的朋友会跟其他两三个人讲述这件事。即使是听到一个秘密，也是听众需要分享的情感事件。

第6章　说谎的语言

1.关于坦白的生理学研究涉及人们将他们生活中最具创伤的经历录进录音机。故事中所有可能具备识别性的信息都已更改。参见彭尼贝克、休

斯和奥赫伦（1987）的文章。

2.对于任何寻求研究自我欺骗相关英文文献的人，我推荐吉姆·马格努森的作品。他在这个和其他课题上的智慧对我来说是非常宝贵的资源。

3.自我欺骗、欺骗和营销之间有一个共通点是“诠释”。在政界，诠释被定义为粉饰真相的艺术。事实上，在重要的辩论或演讲之后，政治工作者经常会约见新闻界的成员，通过诠释让自己候选人的演讲变得更好，同时使对手的演讲听起来更糟。加拿大皇后大学的前沿计算语言学家大卫·斯基利康开发了一种基于欺骗相关的语言特征来识别“诠释”的模型（research.cs.queensu.ca/home/skill/election/election.html）。

4.在《精神分析引论》（Introductory Lectures on Psychoanalysis）中，弗洛伊德花费了大量的笔墨在口误上。

回头想想，你可以理解为什么他通过口误来介绍自己的一般理论。口误每天都很常见，它揭示了人们真正在思考什么。

5.我欠梅兰妮·格林伯格一个人情，感谢她让我再度分析她的语言数据。

6.有关史蒂芬·格拉斯案例的有趣记录可以在www.rickmcgin nis.com/articles/Glassindex.htm 上找到。对于这些分析，我只检查了他41个故事中的39个（一个是与别人共同完成的，另一个完全由已发表的对白组合而成）。

不知为何，与这个讨论相称的一句引用出自玛丽·麦卡锡。1979年，电视台主持人迪克·卡维特采访她时，聊到莉莲·赫尔曼，她以编造麦卡锡的一些故事而闻名。当被问及赫尔曼作品时，麦卡锡哼了一声，说道：“她写的每一个字都是谎言，包括‘和’和‘这’”。赫尔曼随后对麦卡锡提起了诽谤诉讼，但是随着1984年赫尔曼的去世而放弃。

7.还有其他一些引人入胜的文献值得一提，有关使用电脑化文本分析识别现实世界高风险的欺骗。例如，2001年安然由于公司欺诈性会计活动而面临破产，大卫·斯基利康及其同事，马克斯·鲁沃西、冈·肖明及其同事，以及世界各地的其他实验室都研究了安然公司员工在公司破产前十年内的电子邮件。此外，大卫·拉克尔和阿纳斯塔西娅·察可尤基娜通过电脑化文本分析来区分具有欺骗性的首席执行官与真实的首席执行官，文本来自季度收益电话会议。

8.我和丹妮丝·赫德尔的项目还没有提交出版。另外一个发现尤其值得

注意。回想一下，“我”字表示清白。更仔细的分析表明，并不是所有的“我”字都表示清白。被告使用“我”（我会、我是等等）越多，他们越有可能是清白的。事实上，真正有罪的人更多的是用“我”的宾语形式。

9.约会项目是由托马、汉考克和埃里森共同完成（2008年）。

10.差异动词或者说情态动词，区分理想状况和真实状况。“我应该吃蔬菜”，表示我没有吃（现实），但理想中的那个人应该吃。哥伦比亚大学心理学家托里·希金斯围绕自我差异发明了一个详尽的理论，它对目标、动机和心理健康都有影响。理想-现实差异在罗伯特·维克伦德关于自我意识的作品中也有介绍。

11.我喜欢施为句。它们不仅在心理学上很有趣，而且是哲学中使人大开眼界的辩论的核心。参见哲学家约翰·塞尔和约翰·奥斯汀的作品。

12.虽然这里没有讨论，但另一个欺骗的言语特征是使用诸如“嗯”或者“呃”这样的词语。乔安妮·阿尔丘利和同事最近的一篇文章值得一读。另外，请参阅迈克尔·埃拉德的书《嗯》（Um）。

第7章　语言体现地位、权力和领导力

1.回想一下，代词告诉我们人们正在关注什么。许多我们观察到的代词效果与非人类灵长类动物的注视相一致。对于黑猩猩社会等级进行的愉快分析，请参阅法兰斯·德·瓦尔的作品。杰克·多维迪奥和同事们讨论了人类社会中有关统治的视觉线索。

2.约翰·迪恩，个人通讯，2002年8月30日。

3.感谢伊森·伯里斯和他的同事们允许我们再分析他们的数据。

4.有关领导力的文献越来越复杂。要获得一丝方向，请参阅佛瑞德·菲德勒的开创性作品。有关男女领导力最好的研究总结出自爱丽丝·艾格利和她的同事们之手。大卫·瓦尔德曼和他的团队在研究领导力方面做得很好。

5.道格·索弗发现了跨语言和跨文化的语言转变的另一个例子，并作为自己论文的一部分。他发现，1944～1958年，根据社会阶层的不同，寄给南美国家哥伦比亚总统的信中使用第一人称单数的方式不同。你可能会猜到，作者的社会阶层越低，使用“我”字的频率越高。另见豪伊·贾尔斯关于调整理论的开创性工作，这个理论旨在了解人们会在何时转变他们的语言来配合互动伙伴。

6.感谢乔治·西奥多里迪斯对古希腊语的研究（个人通讯，2009年11

月6日）。

7.请注意，我们还发现有其他与地位相关联的语言维度。虽然差别不大，但是具有较低地位的人更倾向于使用以下单词类别：否定词、非人称代词、确定词、脏字。

第8章　爱的语言

1.短信互动是斯拉切和彭尼贝克（2006）项目的一部分。伊丽莎白·哈塞尔贝克和罗茜·奥唐纳之间的直播争斗已经酝酿好几个月了。奥唐纳那时已经宣布将在三个星期内离开这个节目。两位女性之间的整体互动显示的语言风格匹配度（或LSM）高达0.94，这是非常高的分数。

2.对镜像神经元系统的研究仍然是有争议的。越来越多的研究展示了位于布罗卡区的高度专业化的大脑活动，布罗卡区反映了行为模仿。关于镜像神经元方法的主要反对意见是，没有一致的理论或模型解释它如何工作，或者解释它如何与认知活动关联。里佐拉蒂和格雷格里洛（2004），金柏莉·蒙哥马利和她的同事们（2009）以及科茨等人（2001）提供了特别有趣的论文。

3.与撒谎者相关的LSM项目由汉考克、柯里、哥哈和伍德沃思（2008）进行。关于相同数据进行的额外的文本分析报告，请参见杜兰、霍尔、麦卡锡和麦克纳马拉（2010）的文章。

4.虽然多任务处理具有直观的吸引力，但证据显示，即使是完成中等复杂任务，多任务处理也不是一项有效的技术。奥菲尔、纳斯和瓦格纳（2009）发表了一篇文章，讲述了多任务处理的缺点，特别令人信服。

5.快速约会项目来历复杂。得克萨斯农工大学的教授保罗·伊斯特维克于2010年春季访问了我们部门，和我们描述了他和他在西北部的同事伊莱·芬克尔进行的一些快速交友研究。莫莉·艾尔兰对他的演讲很着迷，并问他是否有兴趣将LSM方法应用于快速约会的文本中。几天之内，莫莉的分析得出了一个值得注意的结果：快速约会对话中的LSM是对后期约会的有力预测。然后，莫莉把快速约会的分析结果加到了理查德·斯拉切的短信项目中，她将该论文发给一家顶级杂志，论文很快被接受和发表。最终版本的论文作者有艾尔兰、斯拉切、伊斯特威克、西泽斯、芬克尔和彭尼贝克（2011）。

6.短信项目最初由斯拉切和彭尼贝克（2006）发布，之后，随着数据的再分析，作者加上了艾尔兰、斯拉切等人（2011）。

7.约翰·戈特曼关于人际关系的研究有很多实际的应用，用以实现良好的婚姻。除了约翰的书籍和文章外，《纽约时报》作家塔拉·帕克·帕普还撰写了一本关于婚姻和关系的同类型书，这本书依赖于最近的一些研究。

8.伊丽莎白·巴雷特和罗伯特·布朗宁，西尔维亚·普拉斯和特德·休斯，以及西格蒙德·弗洛伊德和卡尔·荣格的分析是我和莫莉·艾尔兰在2010年发表的一篇文章的一部分。

第9章　通过语言了解团体、公司和社区

1.一些研究追踪了语言使用及其与成功结婚的关系。不足为奇的是，代词的使用，尤其是第二人称代词复数的使用是可靠的预测因素。参见赛德尔及其同事们（2009）的文章，以及瑞秋·西蒙斯、皮特·戈登和黛安·婵布勒斯（2005）的作品。

2.研究夫妻之间代词使用和心力衰竭与代词使用关系的项目由罗哈博夫及其同事们进行。

塞克斯顿和海姆里奇项目仅专注于飞行模拟研究。布莱恩·塞克斯顿后续通过分析坠毁的飞机驾驶舱录音（2010年4月20日，个人通讯）发现，低频率的第二人称代词复数使用和人为错误之间存在联系。另见范世和海姆里奇的作品。

3.比尔·艾克斯是得克萨斯大学阿灵顿分校的社会心理学家，他率先找到研究自然相互作用的有趣方法。在一个典型的实验中，成对的学生访问艾克斯的研究实验室参与对话。当学生完成问卷调查和同意录像后，实验者会尝试开始拍摄，然后“发现”他的相机坏了。实验者离开实验室，声称要去找技术人员。学生们仍然留在实验室里，通常他们开始互相交谈。

他们不知道的是，还有另一台隐藏的相机正在录制他们的互动。然后，学生们被告知有台隐藏的相机，并要求评价他们每分钟的互动。随着谈话的展开，艾克斯能够看到两个人是如何看待对方的。比尔让我们分析了他的一些互动。我强烈推荐他近期的一本书，《陌生人在一个奇怪的实验室》（Strangers in a Strange Lab）。

既然我们在谈论研究人们行为的现实方法，我强烈推荐桑迪·彭特兰和罗兹·皮卡尔德在麻省理工学院媒体实验室的进行的研究。两个人设计了许多方法，可以追踪人们在日常生活中对自己世界的看法和情绪反应。

4.一种考虑第二人称代词复数使用增多的方法是，随着人们交谈的增多，他们彼此更融合。比尔·斯旺和他的同事们一直在进行一些追踪身份融

合的富有想象力项目。例如，让人们更有团体意识可以增加他们为团体战斗和牺牲的可能性。

5.国防工程由我和安德鲁·斯克兰德、依拉·陶兹斯克负责，由美国桑迪亚国家实验室资助。苏珊·奥多姆和斯蒂芬妮·鲁德的实验持续三年对20名专业治疗师进行观察。研究结果发表在苏珊2006年完成的论文中。

6.恐怖袭击后自杀率下降的研究由伊马德·沙利伯及其同事们发表。2004年的马德里地铁爆炸事件之后的语言和心理变化研究结果由我和意弛尔·费尔南德斯、达里奥·派斯发表。沙特哈特曼总结了卡特里娜飓风后新奥尔良居民的书面文字变化。

7.我之前的研究生学生艾米·冈萨雷斯进行了一个复杂的实验，实验中有两组学生，一组学生面对面交流，另一组线上交流。详细信息在冈萨雷斯、汉考克和彭尼贝克（2010）的文章中发表。第二个项目之前描述过，以商学院学生为实验对象，由伊桑·布里斯及其同事们进行。保罗·泰勒及其同事们进行的一些项目是世界级的，上述两个实验研究可以与之并肩。例如，泰勒在英国警方与人质劫持人员之间谈判成功的转录文本中发现了更高的LSM值。

8.克雷格列表项目是一个更大的研究的一部分，重点是估量社区凝聚力。主要团队成员包括辛迪·钟、依拉·陶兹斯克和我。感谢马克·海沃德为我们提供相关的基尼统计数据。

9.词捕捉研究基于我在1990～2010年收集的录音。它们包括现实世界中1162人的自然对话文件。判别分析显示，对5～7个情境进行交叉验证分类，准确度为80%～84%，猜对的概率在16%～20%。

10.我最喜欢的语言地图之一追踪了饮料的通用名称，如“汽水”“苏打”和“可口可乐”。请查看www.popvssoda.com 。

11.社会语言学世界的巨人之一是宾夕法尼亚大学的威廉·拉博夫。拉博夫率先找到方法研究语言使用和口音是如何根据地域和时间的不同而变化的。例如，他早期的一些研究有关在大城市的街区和社区内发现语言差异。后来，他开始关注美国更大范围内的语言差异。很大程度上是由于拉博夫的影响力，宾夕法尼亚大学在社会交往和语言使用的研究方面占据领先地位。宾夕法尼亚大学拥有语言数据联盟，简称LCD（www.ldc.upenn.edu ），它是世界上最大的文本档案馆之一。此外，一位特别有思想的语言学家马克·利伯曼创建了一个非常有影响力的博客网站：语言日志（languagelog.ldc.upenn.edu）。

12.“我相信”这个项目已经向多个方向发展。基于功能词和内容词，我、辛迪·钟和杰森·伦特弗洛一直在绘制详细的美国版语言地图。

13.一种特别热门的文本分析方法可以观察人们在博客、推文或其他通信中是如何使用情感词汇的。虽然情绪分析只关注人们使用的积极情绪和消极情绪词汇，但可以大概了解一下城市、地区或整个国家的幸福感，相关讨论请参阅亚当·克莱默、杰森·伦特弗洛以及在亚历克斯·赖特《纽约时报》上发表的文章。另外，看看埃里克·韦纳一本非常精彩的书《寻找快乐之国》（The Geography of Bliss），讲的是一个人试图弄明白为什么一些国家比其他国家的幸福指数更高。

14.在推断得州高中语言痕迹时，判别分析表明，我们可以将学生正确分类的比例为19%～20%，其中猜对的概率是11%。

第10章　语言侦察

1.将博文与特定作者匹配可以通过多种方式完成。本章中，我们尝试将当下的博客与多年前撰写的博客匹配给同一个作者。这比将同一阶段写的博客匹配给作者要更难。回想一下20个博客使用者的例子，就算我们有每个人连续几天发布的十条博文。我们从每个人的10条博文中挑选一条，然后将挑选出来的博文单独放在一起。目标是通过阅读已知作者身份的其他九条博文，将这20条“孤儿”博文与20位博客使用者进行匹配。我们的电脑在猜测哪个“孤儿”与哪个博客使用者匹配方面做得更好。总体正确率接近58%（其中猜对的概率是5%）。

2.与联邦党人文集相关的文章除了亚岱尔、莫斯特勒和华莱士的之外，请务必阅读帕特里克·尤奥拉（2006）和杰夫·柯林斯及其同事们（2004）的文章。

3.请给我一分钟，我想和地球上特别特别想知道分析联邦党人文集方法的那20个人聊一下。交叉验证方法基于判别分析，假定团体大小相同。最初的功能词分配方法，将所有未知作者的文本都分配给了麦迪逊，这个方法分配的正确率为92.4%，交叉验证分配的正确率为86.4%。两种方法中功能词加标点符号的正确率分别为98.5%和84.8%。基于14个“破绽”词的分析使用二元（在一篇文章中使用或不使用这个词），并且分类和交叉验证准确性都为98.5%。一个分配错误出现在第41章，这出自麦迪逊之手。破绽词分析估计，第49章、第52～57章和第63章的作者是汉密尔顿，第50章、第51章和第62章的作者是麦迪逊。

而汉密尔顿称未知作者的文章中有11篇是他写的，另外三篇由麦迪逊和他共同撰写。麦迪逊后来回忆称，他（麦迪逊）写这些文章时用到汉密

尔顿提供的一些补充意见。所有语言分析表明，共同撰写的文章与汉密尔顿或麦迪逊单独写作的文章完全不同。鉴于此，相比麦迪逊，我倾向于相信汉密尔顿的叙述。

4.特里·佩蒂约翰和唐纳德·萨科（2009）最近的一个项目分析了1955～2003年公告牌单曲榜排名第一的歌曲的歌词。他们发现，在经济衰退期间，人们喜欢更复杂、更社会化和更未来化的歌词。

5.有几种方法可以确定合作导致平均结果还是协同效应。想想约翰·列侬和保罗·麦卡特尼在他们的歌词中是如何使用现在时态动词的。他们单独写的歌曲中，列侬使用现在时态动词的数量一直都高于麦卡特尼（15.8%和13.7%）。根据平均结果假设，他们的合作应该导致现在时态动词比例位于13.7%～15.8%。事实上，列侬-麦卡特尼面对面合作的歌曲中现在时态动词比例为17.6%。在这种情况下，列侬是麦卡特尼和列侬-麦卡特尼之间的平均作者。我们可以从语言学角度计算列侬、麦卡特尼和列侬-麦卡特尼制作的歌曲，哪个处在另外两个中间。统计显示，披头士的平均作者是：列侬50.6%，麦卡特尼36.1%。

列侬-麦卡特尼13.3%。统计显示，联邦党人文集平均作者是：汉密尔顿39.5%，麦迪逊53.9%，汉密尔顿-麦迪逊6.6%。换句话说，列侬-麦卡特尼和汉密尔顿-麦迪逊合作的结果比独立作者的数据更为极端。

6.N-元分析已经被用来描述作者。例如，阿特·格雷泽及其同事们开发了语言行为分类器，用于评估句子的前三个词，以确定正在发出什么类型的句子。（例如，“你在吗？”“这给你！”“你在这。”）他们的言语行为分类器可以用来确定两个交互者的相对状态。

7.考虑语言使用的另一种方法是听总统如何创建自己的故事。丹·麦克亚当斯花了大量时间来分析人们讲述的故事，以便进一步了解他们的性格。他最近的工作是对乔治W.布什的分析，很吸引人。

8.也许总统文件的最佳来源是由加州大学圣芭芭拉分校的格哈德·彼得斯指导的美国总统府项目。彼得斯和他的合作者们正在汇集总统文件使之成为最大档案之一，包括演讲、采访、新闻发布会等等。有关更多信息，请访问www.presidency.ucsb.edu/ 。

9.数据基于人称代词和所有情感词汇使用的标准化分数（z值）的总和。为了使所有数据为正值，将常数3.0添加到所得的z值中。

10.虽然富兰克林·罗斯福的新闻发布会被转录为文字，但它们被大量修改过。罗斯福与新闻界成员达成了协议，所以有大量信息不会被公开。

在社会情感语言方面，他是现代总统中最低的。然而，由于他的语言记录被大幅度修改，所以没有被列入新闻发布会语料库中。

11.博世公司引用了2000年PBS纪录片节目《里根》（Reagen）的记录。www.pbs.org/wgbh/amex/reagan/filmmore/description.html

12.奥巴马少用“我”的案例最初是在马克·利伯曼的语言日志博客中发表：http://languagelog.ldc.upenn.edu/nll/·p＝1651 。“我”字的新闻发布会数据包括35次新闻发布会或会议，时间从奥巴马就职典礼的2009年1月到2010年5月。请注意，语言日志博客的创始人马克·利伯曼发布了奥巴马演讲的分析对比结果。

13.关于奥巴马使用“我”字的引用来自乔治·威尔2009年6月7日在《华盛顿邮报》中的文章，以及斯坦利·菲斯2009年6月7日在《纽约时报》中的文章。

14.越来越多的研究人员试图通过语言分析来确定是否有可能预测恐怖主义、极端主义和暴力行为。现在为美国国土安全部工作的艾里森·史密斯分析了世界各地的暴力和非暴力极端主义团体，发现他们的表达方式是截然不同的。例如，那些最常谈及团体内部关系、权力或统治地位的群体是最有可能从事暴力行为的群体。有趣的是，罗伯特·哈格拉德发现，描述联盟的词汇和描述权力的词汇之间的数量差异，不论在现实世界还是小说中都会预示战争的爆发。

15.澳大利亚东南部的塔斯马尼亚是一个与英国大小相近的岛屿。塔斯马尼亚的第一个（有争议）或者说最著名的探险家之一是亨利·赫利尔。我、詹娜·巴德利和格温妮斯·丹尼尔分析了他的信件和期刊，想要了解有关他如何以及为什么死亡的有趣故事，请参阅我们的文章。

16.对大学生入学申请书的研究有几个注意事项。一般来说，传统的学术标准，如大学理事会分数和高中排名与大学成绩相关，相关程度高于我们的语言参数。

附录　识别和解释功能词的便捷指南

1.感谢克劳德·阐姆托姆提供了大量来自创伤后应激障碍患者的语言文本。

参考文献

Adair,D.1944.The authorship of the disputed Federalist Papers.William and Mary Quarterly,1,97-122.(Available at www.jstor.org/stable/1921883 .)

Adler,N.E.,T.Boyce,M.A.Chesney,S.Cohen,S.Folkman,R.L.Kahn,and S.L.Syme.1994."Socioeconomic status and health:The challenge of the gradient."American Psychologist,49,15-24.

Allport,G.W.1961.Pattern and growth in personality.New York:Holt,Rinehart&Winston.

Arciuli,J.,D.Mallard,&G.Villar.2010."Um,I can tell you're lying":Linguistic markers of deception versus truth-telling in speech.Applied Psycholinguistics,31,397-411.

Argamon,S.,M.Koppel,J.W.Pennebaker,&J.Schler.2007.Mining the blogosphere:Age,gender and the varieties of self-expression.First Monday,12(peer-reviewed journal on the Internet).http://firstmonday.org/issues/issue12_9/argamon/index.html.

———.2009.Automatically profiling the author of an anonymous text.In Communi-cations of the Association for Computing Machinery(CACM),52,119-123.

Austin,J.L.1962.How to do things with words.Cambridge,Mass.:Harvard Univer-sity Press.

Baddeley,J.L,G.R.Daniel,&J.W.Pennebaker.In press.How Henry Hellyer's lan-guage use foretold his suicide.Death Studies.

Balz,Dan,&Haynes Johnson.2009.The battle for America 2008:The story of an extraordinary election.New York:Viking.

Beaver,D.I.2001.Presupposition and assertion in dynamic semantics.Stanford,CA:CSLI Publications.

Beaver,D.,and H.Zeevat 2007.Accommodation.In Oxford Handbook of Linguistic Interfaces,ed.G.Ramchand,and C.Reiss.New York:Oxford

University Press.

Berry,D.S.,J.W.Pennebaker,J.S.Mueller,&W.S.Hiller.1997.Linguistic bases of social perception.Personality and Social Psychology Bulletin,23,526-37.

Biber,D.1988.Variation across speech and writing.Cambridge:Cambridge Univer-sity Press.

Boals,A.,&K.Klein.2005.Word use in emotional narratives about failed romantic relationships and subsequent mental health.Journal of Language and Social Psychology,24,252-68.

Bond,G.D.,&A.Y.Lee.2005.Language of lies in prison:Linguistic classifi cation of prisoners'truthful and deceptive natural language.Applied Cognitive Psychol-ogy,19,313-29.

Borkovic,T.D.2002.Life in the future versus life in the present.Clinical Psychol-ogy:Science and Practice,9,76-80.

Boroditsky,L.2001.Does language shape thought?Mandarin and English speakers'conception of time.Cognitive Psychology,43,1-22.

———.2010.Lost in Translation.Wall Street Journal.July 24.

Bosson,J.K.,W.B.Swann Jr.,&J.W.Pennebaker.2000.Stalking the perfect mea-sure of implicit self-esteem:The blind men and the elephant revisited?Journal of Personality and Social Psychology,79,631-43.

Brandis,W.,&D.Henderson.1970.Social class,language,and communication.London:Routledge.

Brown,R.1968.Words and things:An introduction to language.New York:Free Press.

Burgoon,J.K.,L.A.Stern,&L.Dillman.1995.Interpersonal adaptation:Dyadic interaction patterns.New York:Cambridge University Press.

Burgoon,J.K.,and T.Qin.2006.The dynamic nature of deceptive verbal communi-cation.Journal of Language and Social Psychology,25,76-96.

Burris,E.,M.Rodgers,E.Mannix,M.Hendron,and M.Oldroyd.2009.Playing favorites:The influence of leaders'inner circle on group processes and perfor-mance.Personality and Social Psychology Bulletin,35,1244-57.

Campbell,R.S.,&J.W.Pennebaker.2003.The secret life of pronouns:Flexibility in writing style and physical health.Psychological Science,14,60-65.

Carroll,L.1865.Alice in wonderland.Gutenberg Press,ftp://sunsite.unc.edu/pub/docs/books/gutenberg/etext97/alice30h.zip.

Carstensen,L.L.2009.A long bright future.New York:Random House.

Carver,C.S.,&E.Harmon-Jones.2009.Anger is an approach-related affect:Evi-dence and implications.Psychological Bulletin,135,183-204.

Chen,E.,K.A.Matthews,and W.T.Boyce.2002.Socioeconomic differences in chil-dren's health:How and why do these relationships change with age?Psychological Bulletin,128,295-329.

Chung,C.K.2009.Predicting weight loss in blogs using computerized text analysis.Dissertation Abstracts International:Section B:The Sciences and Engineering,70(9-B),5893.

Chung,C.K.,C.Jones,A.Liu,&J.W.Pennebaker.2008.Predicting success and failure in weight loss blogs through natural language use.Proceedings of the 2008 International Conference on Weblogs and Social Media,180-81.

Chung,C.K.,&J.W.Pennebaker.2007.The psychological functions of function words.In Social communication,ed.K.Fiedler,343-59.New York:Psychology Press.

———.2008.Revealing dimensions of thinking in open-ended self-descriptions:An automated meaning extraction method for natural language.Journal of Re-search in Personality,42,96-132.

———.In press.Linguistic Inquiry and Word Count(LIWC):pronounced"Luke"...and other useful facts.In Applied natural language processing and content analysis:Identification,investigation,and resolution by P.McCarthy and

C.Boonthum.Hershey,Pennsylvania:IGI Global.

Chung,C.K.,P.J.Rentfrow,&J.W.Pennebaker.2011.Geo graph i cal variations in beliefs:Validation of“This I Believe”themes.Talk presented at the 2011 Annual Meeting for the Society of Personality and Social Psychology,San Antonio,TX.

Cialdini,R.B.,R.J.Borden,A.Thorne,M.R.Walker,S.Freeman,&L.R.Sloan.197 in reflected glory:Three(football)fi eld studies.Journal of Person-ality and Social Psychology,34,366-75.

Clark,H.H.1996.Using language.Cambridge,UK:Cambridge University Press.

Clark,H.H.,&S.E.Brennan.1991.Grounding in communication.In Perspectives on socially shared cognition,eds.L.B.Resnick,J.M.Levine,and S.D.Teasley,127-49.Washington,DC:American Psychological Association.

Cohn,M.A.,M.R.Mehl,&J.W.Pennebaker.2004.Linguistic markers of psychologi-cal change surrounding September 11,2001.Psychological Science,15,687-93.

Collins,J.,D.Kaufer,P.Vlachos,B.Butler,&S.Ishizaki.2004.Detecting collabora-tions in text:Comparing the authors’rhetorical language choices in the Federal-ist Papers.Computers and the Humanities,38,15-36.

Damasio,A.R.1995.Descartes’error:Emotion,reason and the human brain.New York:HarperCollins.

Danner,Mark.2005.The secret way to war.New York Review of Books.June 9.www.nybooks.com/articles/archives/2005/jun/09/the-secret-way-to-war/

Davis,D.,and T.C.Brock.1975.Use of first person pronouns as a function of in-creased objective self-awareness and performance feedback.Journal of Experi-mental Social Psychology,11,381-88.

Davison,K.P.,J.W.Pennebaker,&S.S.Dickerson.2000.Who talks?The social psychology of illness support groups.American Psychologist,55,205-17.

De Waal,Frans.2000.Chimpanzee politics,rev.ed.Baltimore:Johns Hopkins Uni-versity Press.

Dickens,C.1843.A Christmas carol.London:Elliot Stock.

Dodds,P.S.,&C.M.Danforth.2009.Measuring the happiness of large-scale written expression:Songs,blogs,and presidents.Journal of Happiness Studies,11,441-56.

Dovidio,J.F.,&S.L.Ellyson.1985.Patterns of visual dominance behaviors in hu-mans.In Power,dominance,and nonverbal behavior,eds.J.F.Dovidio&J.F.Ellyson.New York:Springer-Verlag.

Duran,N.D.,C.Hall,P.M.McCarthy,&D.S.McNamara.2010.The linguistic cor-relates of conversational deception:Comparing natural language pro cessing technologies.Applied Psycholinguistics,31,439-62.

Eagly,A.H.,M.C.Johannesen-Schmidt,&M.L.van Engen.2003.Transformational,transactional,and laissez-faire leadership styles:A meta-analysis comparing women and men.Psychological Bulletin,129,569-91.

Eastwick,P.W.,S.D.Saigal,&E.J.Finkel.In press.Smooth operating:A Structural Analysis of Social Behavior(SASB)perspective on initial romantic encounters.Social Psychological and Personality Science.

Eckert,P.2000.Linguistic variation as social practice.Oxford:Blackwell.

Ekman,P.,M.O'Sullivan,&M.G.Frank.1999.A few can catch a liar.Psychological Science,10,263-66.

Erard,M.2007.Um:Slips,stumbles,and verbal blunders,and what they mean.New York:Random House.

Fast,L.A.,&D.C.Funder.2008.Personality as manifest in word use:Correlations with self-report,acquaintance-report,and behavior.Journal of Personality and Social Psychology,94,334-46.

Fernandez,I.,D.Paez,&J.W.Pennebaker.2009.Comparison of expressive writing after the terrorist attacks of September 11th and March 11th.International Journal of Clinical and Health

Psychology,9,89-103.

Fiedler,F.E.1967.A theory of leadership effectiveness.New York:McGraw-Hill.

Fiedler,K.,&G.R.Semin.1992.Attribution and language as a socio-cognitive envi-ronment.In Language,Interaction,and Social Cognition,eds.G.R.Semin and K.Fiedler,58-78.Thousand Oaks,CA:Sage Publications,Inc.

Finkel,E.J.,P.W.Eastwick,&J.Matthews.2007.Speed-dating as an invaluable tool for studying romantic attraction:A methodological primer.Personal Relation-ships,14,149-166.

Foltz,P.W.1996.Latent semantic analysis for text-based research.Behavior Re-search Methods,Instruments and Computers,28,197-202.

Foster,D.W.2000.Author unknown:On the trail of anonymous.New York:Henry Holt.

Foushee,H.C.,&R.L.Helmreich.1988.Group interaction and flight crew perfor-mance.In Human Factors in Aviation,eds.E.L.Wiener and D.C.Nagel,pp.189-225,San Diego:Academic Press.

Francis,W.N.,&H.Kucera.1982.Frequency analyses of English usage:Lexicon and grammar.Boston:Houghton Miffl in.

Fran.ois,M.,&M.Walker.2010.Towards personality-based user adaptation:Psy-chologically informed stylistic language generation.User Modeling and User-Adapted Interaction,20,227-78.

Fran.ois,M.,M.Walker,M.R.Mehl,&R.Moore.2007.Using linguistic cues for the automatic recognition of personality in conversation and text.Journal of Artifi-cial Intelligence Research,30,457-500.

Frattaroli,J.2006.Experimental disclosure and its moderators:A meta-analysis.Psy-chological Bulletin,132,823-65.

Fredrickson,B.L.2005.The broaden-and-build theory of positive emotions.In The Science of Well-Being,eds.F.A.Huppert,N.Baylis,and B.Keverne,217-38.New York:Oxford University Press.

Fredrickson,B.L.,T.-A.Roberts,S.M.Noll,D.M.Quinn,&J.M.Twenge.1998.That swimsuit becomes you:Sex differences in self-objectification,restrained eating,and math performance.Journal of Personality and Social Psychology,75,269-84.

Freud,S.1901.Psychopathology of everyday life.New York:Basic Books.———.1920/1977.Introductory lectures on psychoanalysis.Translated by James Strachey.New York:Norton.

Galinsky,A.D.,J.C.Magee,D.H.Gruenfeld,K.A.Liljenquist,&J.A.Whitson.200 reduces the press of the situation:implications for creativity,conformity,and dissonance.Journal of Personality and Social Psychology,95,1450-66.

Galinsky,A.D.,J.C.Magee,M.E.Inesi,&D.H.Gruenfeld.2006.Perspectives not taken.Psychological Science,17,1068-74.

Gardner,W.L.,S.Gabriel,&A.Y.Lee.1999."I"value freedom,but"we"value relationships:Self-construal priming mirrors cultural differences in judgment.Psychological Science,10,321-26.

Gilbert,Daniel T.2006.Stumbling on happiness.New York:Alfred A.Knopf.

Giles,H.,J.Coupland,&N.Coupland.1991.Contexts of accommodation:Develop-ments in applied sociolinguistics.New York:Cambridge University Press.

Gill,A.J.,S.Nowson,and J.Oberlander.2009.What are they blogging about?Per-sonality,topic and motivation in blogs.Proceedings of the Third International AAAI Conference on Weblogs and Social Media,Menlo Park,CA,18-25.San Jose,CA.

Gill,A.J.,J.Oberlander,&E.Austin.2006.The perception of e-mail personality at zero-acquaintace.Personality and Individual Differences,40,497-507.

Goffman,E.1981.Forms of talk.Philadelphia:University of Pennsylvania Press.

Gonzales,A.L.,J.T.Hancock,&J.W.Pennebaker.2010.Language

indicators of social dynamics in small groups.Communications Research,37,3-19.

Gortner,E.M.,and J.W.Pennebaker.2003.The archival anatomy of a disaster:Media coverage and community-wide health effects of the Texas A&M Bonfi re Tragedy.Journal of Social and Clinical Psychology,22,580-603.

Gosling,S.D.2008.Snoop:What your stuff says about you.New York:Basic.

Gottman,J.M.,and R.W.Levenson.2002.A two-factor model for predicting when a couple will divorce:Exploratory analyses using 14-year longitudinal data.Family Pro cess,41,83-96.

Gottman,J.M.1995.Why marriages succeed or fail.New York:Simon and Schuster.

Gottschalk,L.A.,&G.C.Gleser.1969.The measurement of psychological states through the content analysis of verbal behavior.Berkeley:University of Califor-nia Press.

Graesser,A.C.,L.Han,M.Jeon,J.Myers,J.Kaltner,Z.Cai,P.McCarthy,L.Shala,M
and classification of speech acts in Arabic discourse.Presented at the nineteenth annual meeting of the Society for Text and Discourse,Rotterdam,the Netherlands,July.

Graesser,A.C.,&S.L.Petschonek.2005.Automated systems that analyze text and discourse:QUAID,Coh-Metrix,and AutoTutor.In Advancing health outcomes research methods and clinical applications,eds.W.R.Lenderking and D.Revicki,McLean,VA:Degnon Associates.

Graesser,A.C.,D.S.McNamara,M.M.Louwerse,&Z.Cai.2004.Coh-Metrix:Analysis of text on cohesion and language.Behavior Research Methods,Instru-ments and Computers,36,193-202.

Graham,H.E.,R.Glaser,T.J.Loving,W.B.Malarkey,J.R.Stowell,&J.K.

Kiecolt-Glaser.2009.Cognitive word use during marital conflict and increases in proinfl ammatory cytokines.Health Psychology,28,621-30.

Graybeal,A.,J.D.Seagal,&J.W.Pennebaker.2002.The role of story-making in disclo-sure writing:The psychometrics of narrative.Psychology and Health,17,571-81.

Greenberg,M.A.,A.A.Stone,&C.B.Wortman.1996.Health and psychological effects of emotional disclosure:A test of the inhibition-confrontation approach.Journal of Personality and Social Psychology,71,588-602.

Groom,C.J.,&J.W.Pennebaker.2005.The language of love:Sex,sexual orientation,and language use in online personal advertisements.Sex Roles,52,447-61.

Halfaker,A.,A.Kittur,R.Kraut,&J.Riedl.2009.A jury of your peers:Quality,experi-ence,and ownership in Wikipedia.Paper presented at ACM Annual Meeting,October 25-27.doi.acm.org/10.1145/1641309.1641332

Hall,J.A.,E.J.Coats,&L.Smith LeBeau.2005.Nonverbal behavior and the vertical dimension of social relations:A meta-analysis.Psychological Bulletin,131,898-924.Hancock,J.T.,D.I.Beaver,C.K.Chung,J.Frazee,J.W.Penneb

Z.Cai.2010.Social Language Processing:A Framework for Analyzing the Com-munication of Terrorists and Authoritarian Regimes.Behavioral Sciences of Terrorism and Political Aggression,2,108-32.

Hancock,J.T.,L.E.Curry,S.Goorha,&M.Woodworth.2008.On lying and being lied to:A linguistic analysis of deception in computer-mediated communication.Discourse Processes,45,1-23.

Hart,B.,&T.Risley.2003.The early catastrophe:The 30 million word gap.Ameri-can Educator,27,4-9.

Hart,R.P.2001.Campaign talk:Why elections are good for us.Princeton,NJ:Princeton University Press.

———.2000.DICTION 5.0:The Text Analysis Program.Thousand Oaks,CA:Sage-Scolari.

Hart,R.P.,S.E.Jarvis,W.P.Jennings,&D.Smith-Howell.2005.Po liti cal key-words:Using language that uses us.New York:Oxford University Press.

Higgins,E.T.1987.Self-discrepancy:A theory relating self and affect.Psychological Review,94,319-40.

Hogenraad,R.2005.What the words of war can tell us about the risk of war.Peace and Conflict:Journal of Peace Psychology,11,no.2,137-51.

Holleran,S.E.,J.Whitehead,T.Schmader,&M.R.Mehl.In press.Talking shop and shooting the breeze:A study of workplace conversation and job disengagement among STEM faculty.Social Psychology and Personality Science.

Hoyt,T.,&E.A.Yeater.2009.Factors affecting women's verbal immediacy to sexu-ally risky situations.Journal of Language and Social Psychology,28,312-19.

Hughes,C.F.1994.Effects of expressing negative and positive emotions and insight on health and adjustment to college.Dissertation Abstracts International:Sec-tion B:The Sciences and Engineering,54,3899.

Ickes,W.2009.Strangers in a strange lab:How personality shapes our initial en-counters with others.New York:Oxford.

Ireland,M.E.,&J.W.Pennebaker.2010.Language style matching in writing:Syn-chrony in essays,correspondence,and poetry.Journal of Personality and Social Psychology,99,549-71.

———.2011.Men imitate life,women underestimate it:Sex differences in script-writers'portrayal of naturalistic dialogue.Manuscript under review.

Ireland,M.E.,R.B.Slatcher,P.W.Eastwick,L.E.Scissors,E.J.Finkel,&J.W.Penn style matching predicts relationship formation and stability.Psychological Science,22,39-44.

James,W.1890.The principles of psychology,vol 1.New York:Henry Holt.

Jamison,K.R.1993.Touched with fire:Manic-depressive illness and the artistic temperament.New York:Free Press.

Josephs,R.A.,J.G.Sellers,M.L.Newman,&P.H.Mehta.2006.The

mismatch effect:When testosterone and status are at odds.Journal of Personality and Social Psychology,90,no.6,999-1013.

Juola,P.2006.Authorship attribution.Foundation and Trends in Informational Re-trieval,1,233-332.

Kacewicz,E.,J.W.Pennebaker,M.Davis,M.Jeon,and A.C.Graesser.Submitted for publication.Pronoun use reflects standings in social hierarchies.

Kahn,J.H.,R.M.Tobin,A.E.Massey,&J.A.Anderson.2007.Measuring emotional expression with the Linguistic Inquiry and Word Count.The American Journal of Psychology,120,263-86.

Kashima,E.,&Y.Kashima.1998.Culture and language:The case of cultural dimen-sions in personal pronoun use.Journal of Cross-Cultural Psychology,29,461-86.

Kaufman,J.C.,&J.D.Sexton.2006.Why doesn't the writing cure help poets?Review of General Psychology,10,268-82.

Kelly,A.2002.The psychology of secrets.New York:Kluwer Academic/Plenum Pub-lishers.

Kelly,A.,&D.Macready.2009.Why disclosing to a confi dante can be good(or bad)for us.In Uncertainty,information management,and disclosure decisions:The-ories and applications,eds.T.D.Afi fi and W.D.Afi fi,(384-402).New York:Routledge/Taylor and Francis Group.

Keltner,D.,D.H.Gruenfeld,&C.Anderson.2003.Power,approach,and inhibition.Psychological Review,110,265-84.

Kim,Y.2008.Effects of expressive writing among bilinguals:Exploring psychological well-being and social behaviour.British Journal of Health Psychology,13,43-47.

King,S.2000.On writing:A memoir of the craft.New York:Scribner.

Kittur,A.,&R.E.Kraut.2010.Beyond Wikipedia:Coordination and conflict in on-line production groups.In CSCW 2010:Proceedings of the ACM Conference on Computer-Supported Cooperative Work,215-24.New York:ACM Press.

Knapp,M.L.,R.P.Hart,&H.S.Dennis.1974.An exploration of deception as a communication construct.Human Communication Research,1,15-29.

Koppel,M.,J.Schler,and K.Zigdon.2005.Determining an author's native language by mining a text for errors(short paper),Proceedings of KDD,Chicago IL,August.

Kotz,S.A.,A.D Ausilio,T.Raettig,C.Begliomini,L.Craighero,M.Fabbri-Destro,

C.Zingales,P.Haggard,&L.Fadiga.2010.Lexicality drives audio-motor trans-formations in Broca's area.Brain and Language,112,3-11.Kramer,A.D.I.2010.An unobtrusive model of"gross national happiness."Proceedings from CHI,2010,287-90.New York:ACM Press.Labov,W.2001.Principles of linguistic change,vol.2:Social factors.Oxford:Black-well.Lakoff,G.,&M.Johnson.1981.Metaphors we live by.Chicago,IL:University of Chicago Press.

Lakoff,R.1975.Language and woman's place.New York:Harper and Row.

Landauer,T.K.,P.W.Foltz,&D.Laham.1998.Introduction to latent semantic analysis.Discourse Processes,25,259-84.

Larcker,D.F.,&A.A.Zakolyukina.2010.Detecting deceptive discussions in confer-ence calls.Stanford GSB Research Paper No.2060;Rock Center for Corporate Governance Working Paper no.83.

Lavergne,Gary M.2010.Before Brown:Heman Marion Sweatt,Thurgood Marshall,and the long road to justice.Austin:University of Texas Press.

———.2006.Bad boy:The true story of Kenneth Allen McDuff,the most notorius serial killer in Texas history.New York:St.Martin's.

Lepore,S.J.,&J.M.Smyth.2002.The writing cure:How expressive writing pro-motes health and emotional well-being.Washington,DC:American Psychologi-cal Association.

Levertov,Denise.1962.The ache of marriage.Poems,1960-1967 by Denise Lever-tov.New York:New Directions.

Little,A.,&D.B.Skillicorn.2010.Patterns of word use for deception detection in testimony.Annals of Information Systems,9,25-40.

Loehlin,J.C.,&N.G.Martin.2001.Age changes in personality traits and their heritabilities during the adult years:Evidence from Australian twin registry samples.Personality and Individual Differences,30,1147-60.

Louwerse,M.M.,K.Lin,A.Drescher,&G.Semin.In press.Linguistic cues predict fraudulent events in a corporate social network.Proceedings of the 32nd Annual Conference of the Cognitive Science Society.

Luria,A.R.1973.The working brain.New York:Penguin.

Lyons,E.J.,M.R.Mehl,&J.W.Pennebaker.2006.Linguistic self-presentation in anorexia:Differences between pro-anorexia and recovering anorexia Internet language use.Journal of Psychosomatic Research,60,253-56.

Maass,A.,M.Karasawa,F.Politi,&S.Suga.2006.Do verbs and adjectives play dif-ferent roles in different cultures?A cross-linguistic analysis of person repre sen-ta tion.Journal of Personality and Social Psychology,90,734-50.

Mairesse,F.,M.A.Walker,M.R.Mehl,&R.K.Moore.2007.Using linguistic cues for the automatic recognition of personality in conversation and text.Journal of Artificial Intelligence Research,30,457-500.

Maner,J.K.,S.L.Miller,N.B.Schmidt,and L.A.Eckel.2008.Submitting to defeat:Social anxiety,dominance threat,and decrements in testosterone.Psychological Science,19,764-68.

Martindale,C.1990.A clockwork muse:The predictability of artistic change.New York:Basic.

McAdams,D.P.2011.George W.Bush and the redemptive dream:A psychological portrait.New York:Oxford University Press.

McCarthy,P.M.,D.Dufty,C.Hempelman,Z.Cai,A.C.Graesser,&D.S.McNa-mara.In press.Evaluating giveness/newness.In Applied natural language processing and content analysis:Identifi cation,investigation,and resolution,eds.P.M.McCarthy and C.Boonthum.Hershey,PA:IGI Global.

McCarthy,P.M.,J.C.Myers,S.W.Briner,A.C.Graesser,&D.S.McNamara.2009. three words all we need?A psychological and computational study of genre recognition.Journal for Language Technology and Computational Linguistics,1,23-57.

McClelland,D.C.1979.Inhibited power motivation and high blood pressure in men.Journal of Abnormal Psychology,88,182-90.

Mehl,M.R.,S.D.Gosling,&J.W.Pennebaker.2006.Personality in its natural habi-tat:Manifestations and implicit folk theories of personality in daily life.Journal of Personality and Social Psychology,90,862-77.

Mehl,M.R.,&J.W.Pennebaker.2003.The social dynamics of a cultural upheaval:Social interactions surrounding September 11,2001.Psychological Science,14,579-85.

———.2003.The sounds of social life:A psychometric analysis of students'daily social environments and conversations.Journal of Personality and Social Psycho-lolgy,84,857-70.

Mehl,M.R.,S.Vazire,N.Ramírez-Esparza,R.B.Slatcher,&J.W.Pennebaker.2007.Are women really more talkative than men?Science,317,82.

Mergenthaler,E.1996.Emotion-abstraction patterns in verbatim protocols:A new way of describing psychotherapeutic processes.Journal of Consulting and Clini-cal Psychology,64,1306-15.

Michel,J-B.,Y.K.Shen,A.P.Aiden,A.Veres,M.K.Gray.the Google Books Team,J.P.Pickett,D.Hoiberg,D.Clancy,J.Norvig,S.Pinker,M.A.Nowak,&E.L.Aide analysis of culture using millions of digitized books.Science,1199644,published online,December 16,2010.

Miller,G.A.1995.The science of words.New York:Scientific American Library.

Montgomery,K.J.,K.R.Seeherman,&J.V.Haxby.2009.The well-tempered social brain.Psychological Science,20,1211-13.

Moore,S.D.,&L.R.Brody.2009.Linguistic predictors of mindfulness in written self-disclosure narratives.Journal of Language and Social Psychology,28,281-96.

Morris,Edmund.1999.Dutch:A memoir of Ronald Reagan.New York:Random House.

———.2005.Conservative compassion.New York Times,August 17,2005.www.nytimes.com/2005/08/17/opinion/17morris.html

Mosteller,F.,&D.L.Wallace.1984.Applied Bayesian and classical inference:The case of the Federalist Papers.New York:Springer-Verlag.

Newman,M.L.,C.J.Groom,L.D.Handelman,&J.W.Pennebaker.2008.Gender differences in language use:An analysis of 14,000 text samples.Discourse Pro-cesses,45,211-46.

Newman,M.L.,J.W.Pennebaker,D.S.Berry,&J.M.Richards.2003.Lying words:Predicting deception from linguistic styles.Personality and Social Psychology Bulletin,29,665-75.

Niederhoffer,K.G.,&J.W.Pennebaker.2002.Linguistic style matching in social interaction.Journal of Language and Social Psychology,21,337-60.

Nisbett,R.E.2003.The geography of thought:How asians and westerners think differently.New York:Free Press.

Nixon,R.M.1974.Submission of recorded presidential conversations to the commit-tee on the judiciary of the house of representatives.Washington,DC:Government Printing Offi ce.

Oberlander,J.,&A.J.Gill.2006.Language with character:A stratifi ed corpus com-parison of individual differences in e-mail communication.Discourse Processes,42,239-70.

Odom,S.D.2006.A qualitative and linguistic analysis of an Authority Issues Training.Doctoral dissertation,University of Texas at Austin.

Ophir,E.,C.I.Nass,&A.D.Wagner.2009.Cognitive control in media multitaskers.Proceedings of the National Academy of Sciences,106,15583-87.

Parker-Pope,T.2010.For better:The science of a good marriage.New York:Dutton.

Pennebaker,J.W.1997.Opening up:The healing power of expressing emotions,rev.ed.New York:Guilford Press.

———.2009.What is“I”saying?The Language Log,guest post,August 9.langua-gelog.ldc.upenn.edu/nll/?p＝1651

Pennebaker,J.W.,R.J.Booth,&M.E.Francis.2007.Linguistic inquiry and word count:LIWC 2007.Austin,TX:LIWC(www.liwc.net).

Pennebaker,J.W.,&C.K.Chung.2008.Computerized text analysis of al-Qaeda transcripts.In A content analysis reader,eds.K.Krippendorff and M.A.Bock,453-65.Thousand Oaks,CA:Sage.

———.In press.Expressive writing and its links to mental and physical health.In Oxford handbook of health psychology,ed.H.S.Friedman,New York:Oxford University Press.

Pennebaker,J.W.,C.K.,Chung,M.Ireland,A.Gonzales,&R.J.Booth.2007.The development and psychometric properties of LIWC2007.Software manual.Austin,TX:LIWC.net.

Pennebaker,J.W.,&A.Gonzales.2008.Making history:Social and psychological pro cesses underlying collective memory.In Collective memory,eds.J.V.Wertsch and P.Boyer,110-29.New York:Cambridge University Press.

Pennebaker,J.W.,C.J.,Groom,D.Loew,&J.Dabbs.2004.Testosterone as a social inhibitor:Two case studies of the effect of testosterone treatment on language.Journal of Abnormal Psychology,113,172-75.

Pennebaker,J.W.,and K.D.Harber.1993.A social stage model of collective coping:The Loma Prieta earthquake and the Persian Gulf War.Journal of Social Issues,49,125-45.

Pennebaker,J.W.,C.F.Hughes,&R.C.O’Heeron.1987.The psychophysiology of confession:Linking inhibitory and psychosomatic processes.Journal of Person-ality and Social Psychology,52,781-93.

Pennebaker,J.W.,&M.Ireland.2008.Analyzing words to understand literature.In New beginnings in literary studies,eds.J.Auracher and W.van Peer,24-48.Newcastle,UK:Cambridge Scholars Publishing.

Pennebaker,J.W.,&L.A.King.1999.Linguistic styles:Language use as an individ-ual difference.Journal of Personality and Social Psychology,77,1296-1312.

Pennebaker,J.W.,&T.C.Lay.2002.Language use and personality during crises:Analyses of Mayor Rudolph Giuliani's press conferences.Journal of Research in Personality,36,271-82.

Pennebaker,J.W.,T.J.Mayne,&M.E.Francis.1997.Linguistic predictors of adaptive bereavement.Journal of Personality and Social Psychology,72,863-71.

Pennebaker,J.W.,M.R.Mehl,&K.G.Niederhoffer.2003.Psychological aspects of natural language use:Our words,our selves.Annual Review of Psychology,54,547-77.

Pennebaker,J.W.,&L.D.Stone.2003.Words of wisdom:Language use over the lifespan.Journal of Personality and Social Psychology,85,291-301.

Pentland,A.2008.Honest signals:How they shape our world.Cambridge,MA:MIT Press.

Perdue,C.W.,J.F.Dovidio,M.B.Gurtman,and Richard B.Tyler.(1990).Us and them:Social categorization and the process of intergroup bias.Journal of Personality and Social Psychology,59,475-86.

Petrie,K.J.,J.W.Pennebaker,&B.Sivertsen.2008.Things we said today:A linguis-tic analysis of the Beatles.Psychology of Aesthetics,Creativity,and the Arts,2,197-202.

Pettijohn II,T.F.,&D.F.Sacco Jr.2009.The language of lyrics:An analysis of pop-u lar Billboard songs across conditions of social and economic threat.Journal of Language and Social Psychology,28,297-311.

Picard,R.W.2010.Emotion research by the people,for the people.Emotion Review,2,250-54.

Pinker,S.1991.Rules of language.Science,253,530-35.

———.1994.The language instinct.New York,NY:William Morrow&Company.

———.2007.The stuff of thought:Language as a window into human nature.New York:Penguin.

Plath,Sylvia.1953.Mad girl's lovesong.In The collected poems of Sylvia Plath,1981.New York:Harper&Row.

Pysczynski,T.,&J.Greenberg.1987.Self-regulatory preservation and the self-focusing style:A self-awareness theory of reactive depression.Psychological Bulletin,102,122-38.

Ramírez-Esparza,N.,C.K.Chung,E.Kacewicz,&J.W.Pennebaker.2008.The psychology of word use in depression forums in English and in Spanish:Testing two text analytic approaches.Proceedings of the 2008 International Conference on Weblogs and Social Media,102-8.

Ramírez-Esparza,N.,C.K.Chung,G.Sierra-Otero,&J.W.Pennebaker.In press.Cross-cultural constructions of self-schemas:Americans and Mexicans.Journal of Cross-Cultural Psychology.

Ramírez-Esparza,N.,S.D.Gosling,V.Benet-Martínez,J.Potter,&J.W.Pennebaker.2006.Do bilinguals have two personalities?A special case of frame switching.Journal of Research in Personality,40,99-120.

Ramírez-Esparza,N.,M.R.Mehl,&J.W.Pennebaker.2009.Are Mexicans more or less sociable than Americans?Insights from a naturalistic observational study.Journal of Research in Personality,43,1-7.

Rentfrow,P.J.,S.D.Gosling,&Potter,J.2008.A theory of the emergence,persis-tence,and expression of geographic variation in psychological characteristics.Perspectives on Psychological Science,3,339-69.

Rentfrow,P.J.,C.Mellander,R.Florida.2009.Happy states of America:A state-level analysis of psychological,economic,and social well-being.Journal of Research in Personality,43,1073-82.

Rimé,B.,D.Paez,P.Kanyangara,&V.Yzerbyt.2011.The social sharing of emotions in interpersonal and in collective situations:Common psychosocial consequences.In Emotion regulation and well-

being,eds.I.Nyklicˇek,A.Vingerhoets,and M.Zeelenberg,147-63.New York:Springer Science＋Business Media.Rizzolatti,G.,&L.Craighero.2004.The mirror-neuron system.Annual Review of Neuroscience,27,169-92.Robins,R.W.,J.L.Tracy,K.Trzesniewski,J.Potter,&S.D.G ity correlates of self-esteem Journal of Research in Personality,35,463-82.

Rohrbaugh,M.J.,M.R.,Mehl,V.Shoham,E.S.Reilly,&G.A.Ewy.2008.Prognos tic significance of spouse“we”talk in couples coping with heart failure.Journal of Consulting and Clinical Psychology,76,781-89.

Rosenberg,S.D.,&G.J.Tucker.1978.Verbal behavior and schizo phrenia:The se-mantic dimension.Archives of General Psychiatry,36,1331-37.

Rude,S.S.,E.M.Gortner,&J.W.Pennebaker.2004.Language use of depressed and depression-vulnerable college students.Cognition and Emotion,18,1121-33.

Salib,E.2003.Effect of 11 September 2001 on suicide and homicide in England and Wales.British Journal of Psychiatry,183,207-12.

Salib,E.,&M.Cortina-Borja.2009.Effect of 7 July 2005 terrorist attacks in London on suicide in England and Wales.British Journal of Psychiatry,194,80-85.

Scholand,A.J.,Y.R.Tausczik,&J.W.Pennebaker.2010.Linguistic analysis of work-place computer-mediated communication.Proceedings of Computer Supported Cooperative Work 2010.

Scott,M.2010.WordSmith Tools v.5.Liverpool:Lexical Analysis Software.

Searle,J.1989.How performatives work.Linguistics and Philosophy,12,535-58.

Seider,B.H.,G.Hirschberger,K.L.Nelson,&R.W.Levenson.2009.We can work it out:Age differences in relational pronouns,physiology,and behavior in mari-tal confl ict.Psychology and Aging,24,604-13.

Seih,Y.,C.K.Chung,&J.W.Pennebaker.In press.Experimental

manipulations of perspective taking and perspective switching in expressive writing.Cognition and Emotion.

Semin,G.R.,&K.Fiedler.1988.The cognitive functions of linguistic categories in describing persons:Social cognition and language.Journal of Personality and Social Psychology,54,558-68.

Semin,G.R.,M.Rubini,&K.Fiedler.1995.The answer is in the question:The ef-fect of verb causality on the locus of explanation.Personality and Social Psy-chology Bulletin,21,834-41.

Sexton,J.B.,&R.L.Helmreich.2000.Analyzing cockpit communications:The links between language,performance,and workload.Human Performance in Ex-treme Environments,5,63-68.

Silver,R.C.,&C.B.Wortman.2007.The stage theory of grief.JAMA:The Journal of the American Medical Association,297,2692.

Simmons,R.A.,D.L.Chambless,&P.C.Gordon.2008.How do hostile and emo-tionally overinvolved relatives view relationships?What relatives'pronoun use tells us.Family Process,47,405-19.

Simmons,R.A.,P.C.Gordon,&D.Chambless.2005.Pronoun use in marital interac-tion:What do"you"and"I"say about marital health?Psychological Science,16,932-36.

Skillicorn,D.B.2008.Knowledge discovery for counterterrorism and law enforcement.Toronto:CRC Press.

Slatcher,R.B.,C.K.Chung,J.W.Pennebaker,&L.D.Stone.2007.Winning words:Individual differences in linguistic style among U.S.presidential and vice presi-dential candidates.Journal of Research in Personality,41,63-75.

Slatcher,R.B.,&J.W.Pennebaker.2006.How do I love thee?Let me count the words:The social effects of expressive writing.Psychological Science,17,660-64.

Smith,A.G.2008.The implicit motives of terrorist groups:How the needs for affi li-ation and power translate into death and destruction.Po liti cal Psychology,29,55-75.

Sofer,D.O.2003.El Pueblo and La Rosca[electronic resource]:A political dialogue in Colombia,1944-58.Austin,Texas:University of Texas Libraries.

Spera,S.P.,E.D.Buhrfeind,&J.W.Pennebaker.1994.Expressive writing and cop-ing with job loss.Academy of Management Journal,37,722-33.

Stirman,S.W.,&J.W.Pennebaker.2001.Word use in the poetry of suicidal and non-suicidal poets.Psychosomatic Medicine,63,517-22.

Stone,L.D.,&J.W.Pennebaker.2002.Trauma in real time:Talking and avoiding online conversations about the death of Princess Diana.Basic and Applied So-cial Psychology,24,172-82.

Stone,P.J.,D.C.Dunphy,&M.S.Smith.1966.The General Inquirer:A Computer Approach to Content Analysis.Cambridge,MA:MIT Press.

Swann Jr.,W.B.,A.Gómez,C.Huici,F.Morales,&J.G.Hixon.2010.Identity fusion and self-sacrifice:Arousal as catalyst of pro-group fighting,dying and helping behavior.Journal of Personality and Social Psychology,99,824-41.

Tannen,D.1990.You just don’t understand.New York:William Morrow&Com-pany.

Tausczik,Y.R.,&J.W.Pennebaker.2009.Leadership in informal groups:A linguistic investigation of Wikipedia.Presented at the annual meeting of Group Processes and Intergroup Relations,Society for Personality and Social Psychology,Las Ve-gas,NV,January.

———.2010.The psychological meaning of words:LIWC and computerized text analysis methods.Journal of Language and Social Psychology,29,24-54.

———.2011.Predicting the quality of online mathematics contributions from users’reputations.Proceedings of the 2011 Conference for the Computer Human Interaction Society(CHI 2011).

Tavris,C.,&E.Aronson.2007.Mistakes were made(but not by me):Why we justify foolish beliefs,bad decisions,and hurtful acts.Orlando,FL:Harcourt.

Taylor,P.J.,&S.Thomas.2008.Linguistic style matching and negotiation outcome.Negotiation and Conflict Management Research,1,263-81.

Toma,C.L.,J.T.,Hancock,&N.B.Ellison.2008.Separating fact from fi ction:An examination of deceptive self-presentation in online dating profi les.Personality and Social Psychology Bulletin,34,1023-36.

Trivers,R.L.2011.Deceipt and self-deception.New York:Basic.

Vano,A.M.2001.Linguistic predictors of treatment success among female sub-stance abusers.Unpublished doctoral dissertation,University of Texas at Aus-tin.

Vano,A.M.,&J.W.Pennebaker,1997.Emotional vocabulary in bilingual Hispanic children:Adjustment and behavioral effects.Journal of Language and Social Psychology,16,191-200.

Vrij,A.,&S.Mann.2006.Criteria-Based Content Analysis:An empirical test of its underlying processes.Psychology,Crime and Law,12,337-49.

Waldman,D.A.,G.G.Ramirez,R.J.House,&P.Puranam.2001.Does leader-ship matter?CEO leadership attributes and profitability under conditions of perceived environmental uncertainty.Academy of Management Journal,44,134-43.

Wegner,D.M.1989.White bears and other unwanted thoughts:Suppression,obses-sion,and the psychology of mental control.New York:Viking/Penguin.

Weiner,E.2008.The geography of bliss:One grump’s search for the happiest places in the world.New York:Twelve.

Weintraub,W.1981.Verbal behavior:Adaptation and psychopathology.New York:Springer.

———.1989.Verbal behavior in everyday life.New York:Springer.

Weston,D.2007.The political brain:The role of emotion in deciding the fate of the nation.New York:Public Affairs.

Wicklund,R.A.1975.Objective self-awareness.Advances in

Experimental Social Psychology,8,233-75.

Wierzbicka,A.1997.Understanding cultures through their key words.New York:Oxford University Press.

Wilson,Timothy D.2002.Strangers to ourselves:Discovering the adaptive uncon-scious.Cambridge,MA:Belknap Press.

Wilson,Timothy D.2011.Redirect:The surprising new science of psychological change.New York:Little,Brown.

Winter,D.G.1998.A motivational analysis of the Clinton first term and the 1996 presidential campaign.The Leadership Quarterly,9,367-76.

Winter,D.G.,&D.C.McClelland.1978.Thematic analysis:An empirically derived measure of the effects of liberal arts education.Journal of Educational Psychol-ogy,70,8-16.

Wolf,M.,C.K.Chung,&H.Kordy.2010.Inpatient treatment to online aftercare:E-mailing themes as a function of therapeutic outcomes.Psychotherapy Research,20,71-85.

———.2010a.MEM's search for meaning:A rejoinder.Psychotherapy Research,20,93-99.

Wolf,M.,A.B.Horn,M.R.Mehl,S.Haug,J.W.Pennebaker,&H.Kordy.2008.Com quantitative Textanalyse:.quivalenz und Robustheit der deutschen Version des Linguistic Inquiry and Word Count[Computer-aided quantitative text analysis:Equivalence and robustness of the German adaption of the Linguistic Inquiry and Word Count].Diagnostica,54,85-98.

Woods,K.M.2007.Iraqi Perspectives Project:Saddam and terrorism:Emerging insights from captured Iraqi documents,vols.1.Institute for Defense Analysis,Alexandria,VA.www.fas.org/irp/eprint/iraqi/index.html

Wright,A.2009.Mining the web for feelings,not facts.New York Times,August 23.www.nytimes.com/2009/08/24/technology/internet/24emotion.html

Yarkoni,T.2010.Personality in 100,000 words:A large-scale analysis

of personality and word use among bloggers.Journal of Research in Personality,44,363-73.

Zhou,L.,J.K.Burgoon,J.F.Nunamaker,&D.Twitchell.2004.Automating linguistics-based cues for detecting deception in text-based asynchronous computer-mediated communications.Group Decision and Negotiation,13,81-106.

Zijlstra,H.,T.van Meerveld,H.van Middendorp,J.W.Pennebaker,&R.Geenen.2004.De Nederlandse versie van de Linguistic Inquiry and Word Count(LIWC),een gecomputeriseerd tekstanalyseprogramma[Dutch version of the Linguistic Inquiry and Word Count(LIWC),a computerized text analysis pro-gram].Gedrag&Gezondheid,32,273-83.